DOCUMENT RESUME

ED 058 858 HE 002 809

AUTHOR	Dressel, Paul L., Ed.
TITLE	The New Colleges: Toward an Appraisal.
INSTITUTION	American Association for Higher Education, Washington, D.C.; American Coll. Testing Program, Iowa City, Iowa.
REPORT NO	Monogr-7
PUB DATE	71
NOTE	327p.
AVAILABLE FROM	The American College Testing Program, P.O. Box 168, Iowa City, Iowa 52240 ($3.00)
EDRS PRICE	MF-$0.65 HC-$13.16
DESCRIPTORS	*Curriculum Evaluation; *Educational Development; *Experimental Colleges; *Higher Education; *Innovation

ABSTRACT

This report contains an anthology of papers presented at the 26th National Conference on Higher Education held by the American Association for Higher Education. The papers present information concerning programs and developments in some of the nation's newer colleges that have grown up as a result of the demand for fewer traditional programs in higher education and more innovation and experimentation. The papers, as a whole, seem to show that it is time for in-depth evaluation and reorganization of these new colleges, for goals seem not to have been clearly defined at the outset of their creation. Many of the colleges have had to retreat to traditional curricula to maintain their existence. Others are struggling along, facing the eminent danger of closing doors with the end of each new semester. It is hoped that the efforts recorded here will encourage those colleges represented in this volume and others in similar situations to reflect deeply on their obligation to engage in evaluation, not only for the benefit of their own programs, but to provide concrete evidence upon which others may reach conclusions as to the desirability of commitment of more resources into these or other innovations. (HS)

MONOGRAPH SEVEN:

THE NEW COLLEGES: TOWARD AN APPRAISAL

[illegible]

ED 058858

"PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY American College Testing Program TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE U.S. OFFICE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER."

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

PUBLISHED BY:

HE002 807

2

ED 058858

THE NEW COLLEGES

ED 028828

CONTRIBUTORS

Alexander W. Astin
Director, Office of Research
American Council on Education

Donald R. Brown
Professor of Psychology
Senior Research Scientist
Center for Research on
Learning and Teaching
University of Michigan

Arthur W. Chickering
Office of Research
American Council on Education

Nancy S. Cole
Research Psychologist
Developmental Research Department
Research and Development Division
The American College Testing Program

Paul L. Dressel
Assistant Provost and Director
of Institutional Research
Michigan State University

John Elmendorf
President
New College

Herbert Garfinkel
Dean, James Madison College
Michigan State University

Charles Harwood
Dean, Fairhaven College
Michigan State University

James A. Kulik
Assistant Professor of Psychology
Research Scientist
Center for Research on
Learning and Teaching
University of Michigan

Larry H. Litten
National Opinion Research Center
University of Chicago

Warren Bryan Martin
Coordinator of Development
Center for Research and Development
in Higher Education
University of California, Berkeley

Beth B. Mason
Director, Counseling Center
Callison College
University of the Pacific

Pressley C. McCoy
Chancellor, Johnston College
University of Redlands

Douglas R. Moore
Provost, Callison College
University of the Pacific

Leo A. Munday
Vice President
Research and Development Division
The American College Testing Program

Theodore M. Newcomb
Associate Director
Residential College
University of Michigan

David J. Reimer
Research Assistant
Center for Research on
Learning and Teaching
University of Michigan

William R. Revelle
Research Assistant
Center for Research on
Learning and Teaching
University of Michigan

Lloyd J. Ring
Executive Assistant
to the Chancellor
University of California at
Santa Cruz

D. Gordon Rohman
Dean, Justin Morrill College
Michigan State University

MONOGRAPH SEVEN:

THE NEW COLLEGES: TOWARD AN APPRAISAL

PAUL L. DRESSEL, Editor

PUBLISHED BY:

THE AMERICAN COLLEGE
TESTING PROGRAM
AND
THE AMERICAN ASSOCIATION
FOR HIGHER EDUCATION

The American College Testing Program is dedicated to the enrichment of education. It was founded as an inviolate public trust and operates as a nonprofit corporation governed by educational representatives from individual states or regions and a Board of Trustees.

A fundamental goal of The Program is to exercise educational leadership by conducting testing, information gathering, evaluating, and related activities in order to (1) assist in the identification and solution of educational problems and (2) communicate to the general and professional publics knowledge and ideas about education.

The chief beneficiaries of The Program's services are students, secondary schools, institutions of higher education, and educational researchers.

Library of Congress Catalog Card Number: 79-172054

©1971 by The American College Testing Program
P.O. Box 168, Iowa City, Iowa 52240

All rights reserved

Printed in the United States of America

Price: $3.00

PREFACE

In the past few years I have been bemused by the inadequacies of universities in meeting the several responsibilities implied by the rubrics of instruction, research, and public service. Various innovations in organization and in governance have been attempted to correct the situation. Research institutes have become profuse, as have other centers, institutes, and programs aimed at analysis and solution of social problems. Faced with the inflexibility, irresponsibility, recalcitrance, and even arrogance of existing departments and colleges which insist on serving their own interests in their own way, administrators, with the assistance of a minority of the faculty, have turned to the development of new units to provide a more relevant and flexible form of undergraduate education. These have been modeled in part on the small liberal arts college and have, in most cases, attempted to emulate the more progressive examples of this type of institution.

I have had some small role in the initiation of such ventures and also the opportunity of visiting many of them. Generally, I have been disheartened by the gap between the ideals as originally planned or as later portrayed in literature and speeches and the reality. Original plans seem to be ignored as such new units develop because each administrator and his faculty tend to mold the program to fit their own prejudices. Without defining the goals of education in terms of expected impacts on students, they rush into designing new courses and become fascinated with new forms of experience and especially group interactions, the rationales for which are unclear and the educational results of which are uncertain. They become quickly as irresponsible and as arrogant as their colleagues in traditional units in insisting that they must run their own show, and that no evaluation is possible except on their terms. New colleges, started as experiments by universities to explore the costs and benefits of alternative forms of undergraduate education, have not only failed to meet this obligation, but actually have rejected it.

In a sense, these new colleges are unfair to the students who enter them. Other than some vague description of requirements and of experiences, the student has no adequate basis for choice of the program, and neither the student nor the faculty has any conception of what benefits in the way of cognitive and affective growth of significance in later life will emerge from the experience. This sounds like a harsh indictment, but one has only to read the reports in this volume to confirm its truth. Of course I speak from an evaluator's viewpoint, and one which some of the contributors to this volume view as an adversary position to their own. Yet my chagrin and irritation lie as much in the anticipation of the effects of these ventures on innovation in undergraduate education as on the lack of critical examination of the programs.

The seeds of decay are already evident in many of the newer colleges. They are having difficulty in attracting students. In some, the programs have already assumed a rigidity which denies freshmen the flexibility which was the original attraction of these programs. Students are disenchanted with the residential aspect in many of the colleges and are moving off campus. Faculty disenchantment is evident in the turnover, the departure to new scenes, and the retreat to more traditional programs. Faculties and administrators not a part of such units are viewing them with increasing distrust and as a strain on scarce resources with no accompanying evidence to justify the continued existence of these atypical programs. And the failure and elimination of such units which will be an inevitable result do not augur well for gaining support for other innovation. And here I speak not as an evaluator, but as one who firmly believes that alternative and successful patterns of undergraduate education must be found.

This, I admit, is an unusual and probably unexpected view for the editor of such a volume to express in a preface. I am not unaware that my colleagues and (former?) friends whom I inveigled into contributing to this volume may regard me as ungrateful and even malicious. So be it. They have spoken frankly of their efforts, their aspirations, and of their difficulties. I only speak equally frankly of my concerns and, most of all, of my disillusionment with the inadequacies of their efforts. But having done so, there remain some other things to be said.

This volume came into being as a result of discussions held with Fred F. Harcleroad, president of The American College Testing Program (ACT), and Leo A. Munday, vice president, Research and Development Division, in connection with their expressed desire to be of assistance in planning the program for the 26th National Conference on Higher Education held by the American Association for Higher Education. We concluded that a program feature on evaluation in the cluster, inner, residential, or new college would be desirable. Out of this agreement and with ACT support this volume emerged. I and AAHE are deeply grateful to Fred Harcleroad and Leo Munday.

I am equally grateful to the contributors who have given many hours of time in preparing their chapters and in appearing at the Conference. Their contributions to this volume present a frank and very revealing picture of the thinking and activities going on in these colleges. The various chapters do reveal that many hours have also been expended by sincere and dedicated persons to exploring new modes of educating youth. They also reveal that there is some awareness of the need for evaluation and perhaps even some feelings of guilt that more has not been done. In several cases there are indications that, belatedly, an adequate evaluation program will be launched. Yet I cannot avoid the suspicion that much of what will be done will be addressed to themselves or, at most, to the group directly involved in such programs rather than to the larger audience of sympathetic observers who are looking for evidence to justify a more extensive remodeling of undergraduate education. Perhaps the best evaluation report here is the one on the University of Michigan college. It tells a great deal about the characteristics of the students and traces some of the changes in attitudes and values which presumably, though not certainly, emerge from the college experience. But it does not tell whether these students have increased in cognitive abilities, in ability to direct and continue their own education, or in ability and commitment to take a responsible adult role in society. I judge that little of the evidence so carefully collected and cogently presented in that chapter will be of much interest to others than like-minded researchers. Certainly I do not find that such evidence convinces adminstrators and faculties generally, or even those directly involved in the innovations, of the merits of such enterprises.

My original hope in promoting this volume was that it would focus attention on the need for evaluation of the new colleges. I think that it clearly demonstrates the need for such evaluation. Litten, Astin, Chickering, and Munday and Cole, in their chapters, have made numerous suggestions of what could and should be done. I hope that the efforts recorded here will encourage those colleges represented in this volume and others in similar situations to reflect deeply on their obligation to engage in evaluation, not only for the benefit of their own programs, but to provide concrete evidence upon which others may reach conclusions as to the desirability of commitment of more resources into these or other innovations.

Finally, I trust that my own highly critical comments will be viewed in their proper light. If even sympathetic observers of this development are somewhat disillusioned, there must certainly be some reason for concern. Innovation in education has too frequently been a leap from one rigidity to another and ultimately equally rigid pattern, for any attempt to achieve complete flexibility leads to chaos and to the imposition of some type of structure. Unless that structure is provided by a statement of objectives so that flexibility in program is always examined and adjusted in relationship to its effectiveness in achieving those goals the inevitable result is a retreat toward traditional patterns and practices. This, as I read it, is the history of innovation in American higher education, and it is particularly likely to be repeated in the universities where resources are slim and traditional practice is intrenched.

Paul L. Dressel

TABLE OF CONTENTS

Chapter 1

The Problems of Evaluation
–Paul L. Dressel 1

Chapter 2

Research for Action
–Arthur W. Chickering 25

Chapter 3

Johnston College: An Experimenting Model
–Pressley C. McCoy 53

Chapter 4

Fairhaven College
–Charles Harwood 89

Chapter 5

The University of Michigan's Residential College
–Theodore M. Newcomb, Donald R. Brown, James A. Kulik,
David J. Reimer, and William R. Revelle 99

Chapter 6

Evaluating Change in New Colleges
–D. Gordon Rohman 143

Chapter 7

Toward a Strategy of Evaluation for a New College
–Beth B. Mason and Douglas R. Moore 155

Chapter 8

New College, Sarasota, Florida
–John Elmendorf 177

Chapter 9

Evaluation for Santa Cruz
–Lloyd J. Ring 185

[continued]

TABLE OF CONTENTS *[continued]*

Chapter 10

James Madison College, Michigan State University
–Herbert Garfinkel 229

Chapter 11

We Know That You're out There, but What Are You Doing?
–Larry H. Litten 257

Chapter 12

Evaluative Research and the New Colleges
–Alexander W. Astin 285

Chapter 13

Evaluation and the New Colleges: Assessment for Student Development
–Leo A. Munday and Nancy S. Cole 301

Chapter 14

Thoughts on Evaluation and Imagination
–Warren Bryan Martin 309

Chapter 1

THE PROBLEMS OF EVALUATION

Paul L. Dressel

Much of what passes as innovation in colleges and universities is really only faddism and tinkering. Changes in requirements, changes in grading practices, freshman seminars, independent study, or alterations in the calendar are often introduced into or grafted onto a program without rethinking the totality of the educational experience and without really modifying the views or the instructional practices of the faculty. Evaluation of the effects of this patchwork approach to innovation is usually sketchy and, in any case, of limited significance because it is impossible to sort out the effects of the patches. They tend to be defended as beneficial rather than being evaluated as a possible element in a coherent total program. Major alterations in the character and impact of higher education are not likely from such patching activity. But the residential, inner, or cluster college holds at least the possibility of a complete reshaping of the undergraduate program by redefining goals, by offering a different environment, and by introducing new content and different methodologies. Yet, though we have had a surfeit of descriptions of such programs, there has been very little in the way of objective evaluation of them. This lack of evaluation is sometimes simply irresponsible, but understandable when the difficulties and hazards of evaluation are understood.

Nature of Evaluation

Evaluation of an educational program is inevitable; the mere attempt to avoid it confirms that an evaluation has already been made. The only issue is the manner in which the evaluation is done and the extent and type of evidence used. When formal, planned evaluation is not present, individuals–administrators, faculty, and students, whether involved in the program or viewing it from the sidelines–make their own. And those who resist formal evaluation cannot reasonably object if the impressionistic evaluations of others are unfavorable.

Concepts of Evaluation

The concepts of evaluation are sufficiently varied that very different approaches can be used. One distinction is that between summative and formative evaluation. Summative evaluation is aimed at appraisal of the overall effectiveness of a program, while formative evaluation is carried on as an aid to development of the program. Summative evaluation, with its implication of finality, assumes a stabilized program about which a decision is to be made. This assumption of stability may be rejected in an experimental (or as some would prefer, an experimenting) program, and certain it is that no administrator or faculty of an innovative program will accept the implications of a negative summative evaluation. Formative evaluation is less threatening, but even a formative evaluation may be distrusted by program proponents who fear that possible negative evidence of a formative evaluation will be interpreted as summative by critics.

Another distinction among approaches to evaluation can be made by conjunction with the prepositions *to, for,* and *with.* Evaluation *done to* implies the collection of evidence followed by summative appraisal by persons external to rather than involved in a program. Traditional faculty grading of students is done *to* them, and evokes from students some of the same fears and frustrations exhibited by faculty members subjected to the same scrutiny. Evaluation *done for* is some improvement in that this may involve a request for the evaluation, and certainly a recognition that the evaluation is necessary, desirable, and ultimately beneficial. When grading or student evaluation of the students by an instructor is used to further experience and improvement, the preposition *for* is applicable. When evaluation is *done with,* it becomes a cooperative endeavor involving students, faculty, and evaluators and has continuing interaction with the program in a *formative* mode. This mode is prone to bias, but is valuable with appropriate safeguards. Evaluation is then a continuing process–an essential component of the program–rather than an occasional and incidental feature.

Viewed in still another manner, evaluation may be regarded as the evaluation of an educational experience with implications for modification, elimination, or extension. It may be regarded also as an educational experience in its own right providing for students' practice in self-appraisal, in clarification and measurement of progress toward goals, and in definition of standards. Finally, evaluation in the sense of making sound judgments based upon evidence and values may be regarded as a goal of education. In this latter sense, the type of evaluation used and its role in the educational process become critical, for it is seen as the principal means for fostering student capability in evaluation and as a model which the students will emulate. It then becomes obvious that the manner in which a faculty of an institution evaluates and reaches decisions about its own activities has a profound effect on students. Students are not likely to become self-directive and mature in appraisal of their own efforts and those of others unless they are involved in evaluation and exposed to an effective model. It is surprising, indeed, that educational innovations which so commonly emphasize student independence in goal determination, setting of standards, and self-evaluation give so little attention to evaluation, either as a basis for study and improvement of the innovations or as an essential experience for the students.

Such evaluation as does exist, like most of the courses provided in our colleges, is past-oriented. Students are tested and graded on how much they know of what has been covered in a course. Evaluation of courses and curriculums–so the faculty insist–must be specific to the content covered and the experiences provided. And if the results are not up to faculty expectation, it is usually the inadequacy of the evaluation and the lethargy or stupidity of the students which are at fault. But evaluation should be future-oriented. The question is not: What does the student know of what has been covered? Rather, it is: What can he do with what he has learned, and what progress has he made toward attainment of competencies accepted as the educational goals? In this latter sense, evaluation becomes the basis for both program revision (when results are generally unsatisfactory for many students) or for planning the further experiences for individuals. Rather than summative appraisal of past program efforts, the focus is on formative appraisal for planning the next steps for curriculum change and for further student experiences. Even at the close of the senior year, this should be true, as is recognized in some illusory sense by our use of the term "commencement."

Patterns of Evaluation

The impact of an educational program is a result of interaction of several factors. There is an input which may be variously viewed as including one or more of the following: students enrolled, faculty, staff, dollars, facilities, and equipment. The educational processes involve the interaction of these inputs in an environment and from this interaction produce some output. The financial benefits conferred on the individual and society could then be related to costs in what is called a cost benefit analysis. Evaluation of educational programs has seldom taken this form; indeed, the tendency is to ignore costs.

Another model views the students (or their characteristics and competencies) as an input exposed to an environment (made up of faculty, staff, facilities, equipment, and educational experiences employing these as well as more subtle affective factors) which fosters change and yields a student output possessing new characteristics and values. This model focuses attention on long-term student change and development. Obviously, measures of relevant student outputs (knowledge, skills, attitudes, values, etc.) are essential to study of college impact in this model. Cognitive outcomes are undoubtedly the most generally accepted goals of students and colleges, but few recent studies have employed these, perhaps because of a combination of the costs, complications, and inadequacies of most such measures. Inputs may be measured in terms of qualities (attitudes, ability, etc.) of individuals which may directly determine outputs or indirectly determine them by interactions with the environment. If inputs can be measured by the same devices used to measure output, the differences yield a measure of gain which has often been used, although it is actually an unreliable measure of change useful, at best, in study of group changes. And even more serious is the fact that the effect of the environment in producing these changes is not readily determinable. Change is surely, in part, a result of college impact, but it is also effected by maturation and out-of-college experiences which are both unknown and uncontrolled. The college environment itself is composed of planned educational experiences and of social-psychological pressures, both of which are affected by the varying characteristics of individuals. Apart from the unique nature of any experience resulting from the differences in individuals, most educational programs provide sufficient flexibility that individual students do not take the same course patterns, have the same instructors, or experience the same associations. Thus, in some sense, an educational program, no matter how well planned, is a highly variable experience for individual students.

Many programs can be regarded as composed of constants and variables; a program may specify certain courses or common experiences which all students are to have but also allow for a great deal of individual adaptation and variability. Any attempt to evaluate such a program which does not in some way attempt to separate out the effects of the constant experiences or at least those which, in reasonably similar form, are experienced by all students and those which vary for individual students will not find the results very helpful in program change. Even though a truly experimental design was used in which there were alternative programs with different mixtures of the differing constant and variable experiences, an evaluation which did not permit the impact to be related to the pattern of experiences would, at most, indicate that a composite program seemed superior to the other without any conclusive evidence as to why. It is not realistic, as occasionally has been suggested, to focus on the constant experiences, arguing that the variable experiences have unknown effects, for it may be the individualized experiences or the interaction of these and the constant experiences that produce the most significant impact.

In some sense, the issue of constants and variables as the focus of attention of evaluation is misleading. Most innovative programs are in some measure a revolt against the static quality of traditional programs. As innovative programs, they go through several years of a dynamic formative development. During this stage, evaluation will inevitably have impact on the program, and thus in a highly dynamic program, evaluation is always formative. On the other hand, when an innovative program has passed the dynamic stage—as some apparently have—and become static, the seeds of discontent and possibly of destruction also have already been planted, although the faculty seldom recognize this. Evaluation at that stage seems to be singularly unproductive in producing change because those who have conducted a "successful" search for truth are not readily brought to renew the search.

Road Blocks to Evaluation

There are many reasons why evaluation is carried out inadequately in most innovative programs. One of the most obvious factors is that of the faculty load and a closely related one is that of the cost. In a new program, faculty members are heavily laden with the task of creating new courses, developing new policies, and, in general, in creating a new educational unit. Much of this is done through very time-consuming committee work. It is commonplace, too, that a limited amount of lead time is provided and faculty members usually find themselves teaching some students while they are still trying to develop the program for the subsequent years. In a very real sense, at this stage, the deliberations themselves are evaluative in nature, subjective though they may be. And, having just beaten out a new course outline or syllabus, there is very little readiness to consider how it might be evaluated or to view it as only part of a total experience, the impact of which must be evaluated.

In fact, many innovative programs are developed gradually over several years. Thus, the nature of the program grows out of the cumulative decisions made by different groups of faculty members over a span of three or four years rather than taking a sequential and unified form based on a commitment to some agreed-upon objectives. The faculty either has not had the time, or has simply refused to take the time, to project the program into the future and consider the evaluation ultimately required to determine effectiveness.

If additional funds were available to cover the cost of an evaluation staff, more evaluation would be done. Even so, the evaluation staff may find itself beset secondhand with the same difficulties that have been noted for the faculty. Evaluators cannot readily move in and evaluate a program for which objectives have not yet been fully formulated. And the attempt to define objectives and develop a formative evaluation program may only mean that the evaluators are co-opted into program development. Thus endeth their evaluation.

Emotional Reactions

Another major roadblock to evaluation is found in the emotional reactions of the people involved. One of the most apparent is the personal involvement of the dean or the administrator who is given the task of developing the program. When the first dean is appointed for a new program, or the first president or chancellor is named for a new institution, this individual finds himself in an unusual situation. He is able to create *on paper* his own ideal of what an educational program should be. He shortly finds that the successive arrival of other administrators, librarians, staff members, and students modifies or chops away aspects of his original idea but, in some sense, he still visualizes this developing new program as his own. His language betrays this, as do his emotional reactions when doubts are expressed or criticisms are made. Most deans or administrators in innovative programs find the attempt to evaluate somewhat distasteful and, in some not too subtle way, a questioning of their own judgment. The faculty and even the students shortly begin to take on some of these defensive mannerisms. An evaluator may see himself as evaluating the program, but the individuals operating that program may see the evaluation as a questioning of their competency. In fact, the people and program are closely related, and unless the faculty, administration, and the students in a program participate in a collection of evidence and weigh the findings about the program's effectiveness, it is not likely that the evaluation will have any impact. The classic remark of one administrator made some years ago comes to mind. "This evaluation will be a waste of time, for it either will demonstrate that the program is excellent or that it is defective in some sense. In the first case, it is a waste of time because we already know that it's a good program and, in the second, it's a waste of time because we would not believe any evidence indicative of weakness."

Students enrolled in new programs seem primarily to value the close associations and interactions with other students and the faculty. They are little inclined to think about long-term objectives and results. To some the immediate experience is so satisfying they refuse to look beyond it. Some resent, like the faculty, any insistence that the worth of experience must be evaluated by results. A few object to the idea that education should change them; rather, it must be pleasant, challenging, and permit student participation and determination of both experiences and results.

Innovators, too, take on a composite character made up of missionary zeal for converting their associates and others who will listen to them, radical tendencies evidenced by their criticism of those programs which they have rejected or are trying to replace, and something of a martyr complex resulting from the criticisms of their former associates and the doubts continually expressed by others that the new program is really as effective as its proponents think it to be. Innovators are both offensive and defensive, and they may become more offensive as they become more defensive. Out of this emerges a natural tendency to resent evaluation reflected in a question: "Why us?" The implication is that other programs around are really not as good as theirs and that traditional programs ought to be evaluated at least simultaneously with if not before the innovative program. In a sense, the "why us" question indicates a conviction that the program is not only good but is actually better than the traditional programs and hence really needs no evaluation. There is also often an implication that no one other than those who are operating the program really understands what it is all about. Accordingly, no one else is capable of evaluation. It is true that traditional and innovative may differ in objectives as well as in the educational experience, but there are generally some common outcomes which would justify comparative evaluations to the improvement of both.

An amusing distinction, too, emerges with regard to the characterization of innovative programs as experimental. The designation of a new program as experimental usually turns out to have unpleasant overtones to those involved in it. It may suggest a rigidly planned pattern and it does suggest an untried one. Students reject the idea of being guinea pigs; faculty reject the idea that their rationally determined conclusions about education need to be validated; and administrators resent the possibility that, having been characterized as an experimental program, the program itself might be eliminated after a period of time because the experiment has failed. Accordingly, preference has been expressed for replacing experimental by experimenting, thereby suggesting a continuing unit with a charge for continuing experimentation. This distinction changes the concept from that of a static new program experimental in nature to a dynamic experimenting program which is subject to constant change and which may provide feedback for the improvement of other programs in an institution. In fact, the distinction does not seem to be a very significant one except in the minds of those who are involved in such programs and therefore highly sensitive to criticism which, to them, always seems to be involved in a demand for evaluation. An experimenting program may also be regarded as an experiment, and one would hope that an experimental program is engaged in experimenting. In actuality, the characterization as either experimental or experimenting is largely erroneous when applied to innovative educational programs. There is nothing approaching the rigor of the scientific experiment and there is, in general, a rejection of the idea that an array of alternatives might be examined and compared. Differences in views rather than yielding alternative programs for experimentation are resolved by compromise or synthesis; and concerns about relevance, however subjectively determined, take precedence over rigor. Many of the decisions in new programs are made on grounds of social and moral passion rather than on either rationality or evidence.

Priorities. In developing a new program, some limited attention may be given to desired outcomes, but major attention seems to be given to the input, to the environment, and to processes carried on in that environment. Usually there will be many discussions of the kinds of students desired, although, in fact, an innovative program may find itself accepting almost all applicants in order to meet its quota. There are also likely to be many discussions of the kind of faculty desired but, here again, the actual search may have to be resolved by taking those who will accept an offer. Much attention has been given to creating a social environment which fosters the affective development of the student while retaining some intellectual character. But perhaps most of all a faculty, in developing a new program, spends most of its efforts on the development of new courses and patterns of instruction. Even the reiterated emphasis on the importance of individual study is often shoved to one side by the gradual expansion of the more structured forms of educational experience. Their courses are, to the faculty, so obviously important that there is little concern about formulation of objectives.

The priority given to the input, to the environment, and to the processes makes it difficult to plan evaluation. Basically, most administrators and faculty members in higher education, even in innovative programs, are primarily interested in cognitive outcomes. Yet the effect of the social environment which is created in a new program is perhaps primarily on affective learning rather than on cognitive achievement. In the area of affective learning an apparently identical experience may be very different for each individual. To a large extent, too, the individual's expectation may determine the results–a self-fulfilling prophecy. Thus, the uncertain relation of program to affective outcomes, as well as the difficulty in defining affective outcomes, pose problems for the evaluator. Perhaps even more troublesome is the tendency, encouraged by social psychologists, to focus on affective concerns to the exclusion of the cognitive outcomes.

Another difficulty in connection with attention to educational processes is found in the fact that the things commonly believed by faculty to promote learning have not been shown to have any significant impact on student learning and attitudes. The phraseology here is important. One should not say that teaching methods, class size, televised or traditional instruction, team and ordinary teaching, types of facilities, etc., have no distinctive impacts but, in the many studies that have been carried on, no significant difference in effects on student learning and attitudes among these different approaches has been demonstrated. Most evaluators know this; most faculty members simply do not believe it. They tend, then, to put great emphasis on some of these factors, particularly small classes, resist any expression of doubts about the need of justification for this, and yet have built into a program certain educational processes whose nature often is not defined and whose impact is imagined rather than real. For example, emphasis on independent study may readily go awry. One investigator, when encouraged to hunt up a faculty member engaged in conference with an independent study student, found the student seated on the front row of a 30-chair classroom with the professor lecturing.

Confusion over goals. One of the major difficulties in any program evaluation is confusion, disagreement, or lack of clarity of the goals of the program. There is a tendency to let educational goals emerge vaguely out of program decisions, though the program development may be preceded by exhortations in the direction of esoteric matters, such as creativity, critical thinking, character, citizenship, personality, and the like. There is, too, a difficulty in agreeing as to who decides the goals of an educational program. Some students, supported by faculty and administrators, will argue that any statement of goals emerging from the faculty is unacceptable. The faculty readily reject goals proposed by administrators, but seldom write any sensible ones of their own. There are many who reject the need for statements of goals, believing them to be obvious or implicit in the courses and experiences provided. Yet there is an obligation to set some goals in an educational program and to inform students and faculty as to what these goals are. Much of the controversy revolves around the specificity of objectives and the confusion between experiences and objectives. Those who demand a precise detailed description of student behavior tend to develop a long list of specific objectives and, in the process, lose sight of some of the ideals, values, and social skills which are important (though somewhat esoteric and intangible) as well as the higher-order cognitive processes, such as analysis, synthesis, and critical or creative thinking. The latter, in the long run, are the most significant outcomes of higher education. Unless these broader outcomes and more pervasive ones are kept in mind, detailed objectives carried to the level of spelling out particular experiences and highly specific knowledge and skill are as destructive to the unity of a program as the traditional emphasis on courses and disciplines. Another element in the controversy about objectives is the distinction between the actual impact and the intended impact. Those who state objectives a priori state their expectations or desires and tend to evaluate in terms of the attainment of these. They recognize that there may be unintended results and seek to learn about these. Others will argue, however, that, having set up the best possible program in terms of experience and judgment, the impacts on individual students are inevitably varied, and that evaluation should focus on determining what those impacts are rather than trying to look at a particular subset agreed upon in advance.

The very range of objectives which has already been touched upon also makes for some difficulty in getting agreement. All too frequently one finds that faculty members, despite everything that may be said about more pervasive concerns, retreat to an emphasis on factual and conceptual mastery of subject matter in their particular courses. The more significant cognitive abilities, such as understanding, analysis, and the like, are then viewed as important but not approachable directly in a class. Rather, the tendency is to assume that these gradually develop in individuals after they have been subjected to a series of well-planned experiences and have mastered enough subject matter. There are those, of course, particularly in loosely structured disciplines which are not cumulative in nature and which have little relevance for many students, who argue that affective learning, values, interests, attitudes, and the like are the most important outcomes of education. Not infrequently, this line of argument ultimately focuses on the interaction of students with specific professors as the main basis for encouraging the development of certain attitudes and values. In this circumstance, any evaluation is fated to irritate the professors. There are also broad intellectual objectives having to do with the range of problems that an individual recognizes, his awareness and concern about them, and related personal and social concerns and activities and obligations. Finally, there are psychomotor objectives which are not only important in the areas of music and art, but also in many professional fields, science laboratories, and the like. While these tend to be recognized in professional programs, they are often largely ignored in undergraduate programs with a liberal education orientation.

Another source of confusion about goals is found in the lack of unity and coherence in a program. Though innovative programs usually rise in part out of the fragmented nature of the undergraduate educational experience, the development of the innovative program does not move far before subdivisions of knowledge are called upon to provide a structural organization. These may be in terms of courses, they may be in terms of divisions, and, more rarely but still possibly, even a departmental or disciplinary organization. Although this may be necessary, the inevitable result is that the efforts of the faculty are devoted to developing their segment of the program; and the threads which tie these elements together at any point in the student's experience or the trails which he follows as he moves from year to year through the experience are obscured. The various groupings of experience and the faculty involved in them tend to develop distinctive conceptions of objectives, thus contributing further to the lack of program coherence.

Still another problem in evaluation is found in relating processes or educational experiences and goals. If evaluation cannot provide feedback as to the significance of particular phases of a program in which particular faculty members have a major interest, then the evaluation is not really very meaningful to them. On the other hand, faculty members, as they plan experiences, do not consciously select these in reference to goals. For example, most faculty members prefer small classes over large ones, and they talk glibly about the importance of discussion even though they may not permit any great amount of it. If one agrees in advance that students should achieve the ability to work together in groups of various sizes in analyzing problems and developing patterns of cooperative endeavor in gaining information about the problems in developing prospective solutions to them, then a small-group experience may be seen as an essential learning experience definitely related to the acquiring of a certain competency. Furthermore, the specification of that competency says something fairly definite about the precise nature of that group learning experience. But, unless this kind of thinking has been indulged in ahead of time, it is probable that the particular educational processes will have been selected primarily on the basis of faculty preference or bias and may be conducted in a manner which has no obvious relationship to any significant educational outcome.

Still another factor to be dealt with is the fact that changing social conditions are reflected by curriculum changes. The traditional organization of knowledge and the rigor demanded by the disciplinarian are constantly being attacked by the demands of society and of the student for more obvious relevance between what is studied and the changing social conditions. The classical curriculum was modified on these grounds. The land-grant university came into being because of this demand. The demands in the current scene are very similar to these. It is not, then, appropriate to take the point of view that some persons have espoused–that is, that education rightly conceived is the same at all times and places. Yet, in some sense, there exist certain competencies in the area of communication–the ability to carry on an independent investigation, and the like–which are always essential outcomes of a higher education experience. If these are not kept in view, the program does degenerate into a set of uncoordinated experiences in which the desires and the concept of relevance of each individual are the governing factors; and, carried to the extreme, the satisfaction with his experience is perhaps the ultimate criterion of the effectiveness of such a program. Momentary satisfaction is hardly an adequate criterion. Indeed, it may be stultifying and injurious to further development. Dissatisfaction, properly focused, may, in many ways, be more conducive to further progress.

Conceptions of teaching and of learning vary a great deal among students and among faculty members. These are reflected in the immediately prior paragraphs which discuss the problem of lack of clarity of goals. Even faculty members who are attracted to innovative programs are often convinced that it is largely the activity of the teacher that determines what the student learns. On the other hand, there is every basis for believing that the effectiveness of a teacher is seen primarily in the extent to which he is able to create enthusiasm in a student for his own learning. Learning is, after all, an individual and personal experience, although teachers can have profound effect on individuals by motivating them, by stimulating them, and providing comments which send the individual back to his studies with renewed enthusiasm. The conception of teaching, however, which emphasizes what the teacher does in the classroom situation tends ultimately to yield an evaluation based upon what the student has learned of what the teacher covered. Differences in the conception of teaching and learning are especially apparent across the cumulative and noncumulative fields of knowledge. Many courses in literature and history, for example, degenerate into the coverage of a highly specific set of factual materials which the students are required to master in order to pass the course. As one moves from one course to another of this type, there is very little in the way of sequential development, little that can be identified as a cumulative educational experience resulting in pervasive or transcendental goals either affective or cognitive. In a cumulative field, such as mathematics or science, it is expected, and indeed it is essential, that, as the student moves through a sequentially related series of courses, he improves to the point of being able to handle certain tasks that he could not have done before and he gets some satisfaction out of the obvious progress that is made in this regard. It is equally true that, in these fields, too, the emphasis can be on evaluating how well the student has learned to do the specific things which have been presented in courses rather than on how much power he has developed to deal with situations and problems which have not been specifically covered. Only as goals or competencies are clearly stated and evaluation viewed as determination of progress toward these can evaluation make a significant contribution to both program and students.

Inadequacy of Evaluation Models

The evaluation model developed by Ralph Tyler, which calls for a definition of objectives, the specification of an appropriate program of experiences, and then an evaluation which can serve either as a feedback for program improvement or as a basis for judging the effectiveness of the program in comparison with other programs, is currently regarded by many persons as an inadequate model. The objections are that faculty members find it peculiarly difficult to specify objectives in any meaningful terms, and that they do not generally undertake to plan the experiences provided in direct relationship to any set of agreed-upon objectives. It may be argued, too, that no definite relationship has been shown between various patterns of educational experiences and the results, although this is often simply because the objectives have been spelled out in such terms that they have no direct implication for program development, and hence have no identifiable relationship to whatever changes do take place in students. The model may also be criticized–although this is not an inherent weakness–for not taking into account explicitly the need to look at the program input and relate the output to this as well as to the experiences provided.

A second model, which may be called the impact model, simply undertakes to determine the kinds of impact which the program has on individuals and on a group of individuals. The impact model, as often used, does not do anything more than demonstrate that some kinds of changes have taken place over a period of time. The effects of maturity and unrelated experiences are not easily separated from those of the planned program. The question of what changes have taken place in what kinds of students and the relation of these changes to experiences and institutional program is a very difficult one to answer.

Some studies have concentrated on demonstrating changes in such affective and intellectually related factors as authoritarianism, liberalism, openmindedness, independence, self-confidence, and impulse expression and have perhaps related these to various curricular groups and institutional characteristics, climates, or environments. The latter have been determined either objectively by a combination of demographic statistics, such as dollar support per student, faculty-student ratio, library holdings, etc., or by profiles prepared by use of various scales, such as CUES. The results have not generated conclusive evidence that these characteristics have definitive impacts on students.

Another model starts with the assumption that the program is so far individualized that one can evaluate only through case studies which follow an individual day by day and cumulatively examine the impact of his experiences on him. While such case studies are revelatory and always useful, it is very difficult to maintain any high degree of objectivity in case studies, particularly when these are carried on by individuals who are themselves emotionally involved in the program. On the other hand, observers from outside carrying on case studies may have biases and may lack adequate understanding of what is being done in order to carry on a case study. Furthermore, there is always the problem that a case study may generate such an intimate relationship between the individual studied and the person carrying on the study that this in itself becomes a significant educational experience, having an impact which is at least as important as some of the planned educational experiences. Finally, when case studies are made, an individual, whether external or internal originally, is likely to become so vitally involved with the individual with whom he works that the development of the individual regarded in a sympathetic way takes precedence over the determination of the impact of the educational program upon that individual.

The course model is one of the more common patterns and fits into the professorial mode of thinking. Whether in a traditional program or in an innovative one, the professor tends to think in terms of a course which he gives. He would like, if he is interested in evaluation at all, to have a report which has direct relevance to what he is doing in that course. Yet the sum total of an educational experience is probably more than the sum of the courses—or at least it should be—and, since students do take different courses, it is not easy in an evaluation of a program to take into account the different course patterns or to relate back to any particular course the educational impact of that course. Evaluating course by course does not add up to an adequate evaluation of the whole program.

Still another model, which has been suggested earlier, is that of the professor. This point of view argues that the significance of an educational program largely depends upon bringing students together with group teachers and that there is an interaction and a character development and motivation which proceed out of this which transcend any other aspect of the program. This might, to some extent, be an hypothesis which is subject to study, but if it has been studied in any objective manner, the results have not been widely reported. The experience of evaluators generally is that, when this point of view is taken, the task of the evaluator becomes very difficult indeed, because the results of evaluation have a highly personal implication which is resented and rejected unless it is laudatory.

The issue of whether a residential college imbedded in a large university can create a distinctive subenvironment is worthy of study. The further question of whether this subenvironment can be demonstrated to produce a greater or a different pattern of student change requires examination. But unless the change is related to input characteristics and to costs, the results are not very useful for a decision on the worth of such colleges.

There are many incidental matters which can be studied as part of an evaluation. A program which demonstrates that it meets the needs of a group of students not otherwise in college or satisfied with college experiences has demonstrated significant merit. A program which shows a higher retention rate than other programs may justify its existence and some additional costs on that basis alone. Student and faculty satisfaction is a desirable incidental result but hardly definitive. Graduate school attendance and success in later life are worthy of study, but probably only informative by comparison with other programs, and not really instructive as to the unique educational benefits of the innovative program.

Perhaps the major inadequacy of all evaluation is that few people, least of all those involved in an educational program, really accept the idea that major changes in or the elimination of a program will follow consequent upon an evaluation of it. The major changes in an educational program have been made on the basis of response to demands or criticisms of society or on the basis of the enthusiasm or ambition of faculty members or administrators. No model yet devised, even with more sensitive instruments and greater unanimity on objectives, can demonstrate to all concerned that one approach to education is superior to another. When it is accepted that education at all levels is broken down into a series of reasonably discrete experiences dominated by different individuals, it seems obvious that no evaluation is ever going to meet the particular concerns of all of those involved in a new program. At the same time, the reasonable criticisms of the inadequacy of evaluation play into the hands of those who would avoid evaluation or who would deny the use of evaluation data as a basis for eliminating a program or demanding major changes in it. Some kinds of evaluation do lead to program elimination. A number of cluster colleges have disappeared simply because the type of program generated in the college proved attractive to an insufficient number of students. Surely the opposite of this must be avoided. A program which attracts and holds large numbers of students is not necessarily a good or successful educational program. But we shall not ascertain this without evaluation.

Methodology[1]

Interrelated with the matter of models is the matter of the methodology which is used. The professional evaluator tends to prefer an experimental mode and makes heavy use of statistics and objective measurements. He seeks more and more rigorous research and more and more precise measurement in the hope that ultimately educational research can answer some basic questions about learning and the relationship of various types of learning experiences in such a way as to improve the quality of education, while always monitoring the costs so as to maximize a cost benefit or cost effectiveness ratio. Educational research, as carried on over a period of years, has not brought us to that promised land as yet. More recent endeavors, which have involved placing large sums of money into particular centers to carry on continuous research in depth, have not yet justified the confidence that many people had that, given this kind of opportunity, educational research would demonstrate its worth. On the whole, in the context of a college, whether an inner or cluster college in a larger institution or a free standing institution, it seems unlikely that the demands of the rigorous experimenter, statistical measurement oriented evaluator can be satisfied. One deals with convenience samples rather than with random samples, with quasi-experiments rather than with real experiments, and when attempts are made to set up matched groups, the adjective "matched" is advisedly placed in quotes.[2]

---

[1] Kenneth A. Feldman has an extensive discussion of this and related issues in "Research Strategies in Studying College Impact," *ACT Research Report No. 34* (Iowa City, Iowa: The American College Testing Program, 1970).

[2] Donald T. Campbell and Julian C. Stanley, "Experimental and Quasi-Experimental Designs for Research on Teaching," Chapter 5 in *Handbook of Research on Teaching*, N. L. Gage, Ed. (Chicago: Rand McNally & Company, 1963).

Attempts to improve this situation by moving to evaluation on an interinstitutional basis are not likely to improve the methodological problem. At the interinstitutional level, difficulties arise in agreeing upon objectives, differences in the experiences provided raise doubts in the minds of individuals about the appropriateness of a particular device or evaluation procedure to their situation, and the tendency is for the move toward evaluation to be slowed by the extensive committee work involved in trying to soothe and reassure alarmed faculty and in trying to arrive at statements of objectives and procedures satisfactory to all concerned. In the interinstitutional process, the focus turns from the evaluation of individual programs to the search for generalities about types of programs and experiences, and precautions are taken so that each program will be adequately safeguarded from criticism by the way in which the evaluation program is developed and carried out.

Some evaluation projects have been content to use the final status of students on the basis of certain measurements without taking into account where the students stood when they entered the program. In a course model of evaluation, this may make some sense with certain courses. Most students who enter a statistics course know little or no statistics prior to the course, and an assessment at the end of the course may be a reasonably good indication of how much they have learned about statistics. On the other hand, any attempt to evaluate the methodology of that course as having produced the change is limited by the fact that no evidence is available on the competencies of the students as they entered the course. Even though they knew no statistics, they might differ significantly in mathematical background or in general aptitude from students who might attempt to take that course at another time or place. Thus an adjusted final status is seen as more valid by most evaluators, but questions do arise as to what variables are appropriate in making adjustments of the final status. If one extends the concept of adjusted final status to try to allow for the multiple tracks followed by various students, the entire situation can become very complicated, and in a small college the numbers of students involved in each subgroup so small as to defy interpretation. And, of course, there are those who would prefer to use some measure of growth from the beginning of an educational experience to the end. The difficulty with this approach, as many analyses have shown, is that the raw score differences between initial and final testing are unreliable, and they have inherent tendencies which pretty much invalidate their use as a satisfactory measure of the impact of an educational program.

Finally, there is a question as to whether the evaluation of a program should be based upon a comparative or an absolute effectiveness in attaining the objectives. Institutions which find that comparison with other institutions makes them look good find this a very satisfactory experience. Those who learn that the performance of their school or of their students is less adequate than that in other institutions are inclined to question the appropriateness of the institutions with which the comparison is made, express doubts of the validity of the instruments used for their particular program, condemn somebody else for the quality of the students that they have to deal with, or possibly point out that the other institutions require, let us say, three more credits in the humanities and that, accordingly, the whole matter can be resolved by adding another requirement.

Measures of absolute effectiveness are hard to come by. Some years ago, teaching a beginning statistics course, I gave the class a list of specific things that they must be able to do to pass the course. These included such items as computing the mean for grouped and ungrouped data, determining the median, computing the standard deviation, and computing correlation. This required further clarification because there are various ways of doing this, and students immediately demanded to know whether they could use any method. There was also an immediate demand to know whether arithmetical errors would be ignored, even though it was pointed out that incorrect statistical calculations are useless and that some of them—for example, a correlation greater than one—are absurdities. Still others wanted to know whether they had to recall the formulas or whether these would be given to them. And, finally, when all of these matters had been clarified, there were still several students who, on the final examination, were unable to carry out these simple computations or confused a mean and a median. Despite the fact that these students answered other questions on the examination, they were failed in accordance with the original insistence upon meeting an absolute standard. The repercussions were such that never again was this approach used. It seems unlikely that measures of absolute effectiveness can be agreed upon, and yet any program which can make judgments only by comparison with some other program can hardly be regarded as having spent sufficient time in defining educational outcomes or in making these sufficiently specific so that an effective educational program can be developed. Surely there are some minimal things that can be demanded in the way of communication skills, in the way of using resources, in gaining understanding of a problem, in the way of detecting the difference between cognitive and affective elements in a discussion. If a college degree, or if completion of a particular educational program, is to have any significance, then there must be some agreement, not only on some common experiences, but also on some accomplishments. Innovative programs should give serious attention to this approach.

Evaluation in the Long Haul

Much of the evaluation which is carried on is done on a short-term basis. There is a desire to know what happens to students after one year, after two years, and possibly after four years. Seldom do we raise the question as to what difference a program really makes in a person's life in contrast with those experiencing other programs. Within a great range of program differences, it is highly probable that the long-term differences are negligible or indiscernible. When new programs are introduced which are significantly more expensive than existing ones, a particular obligation is involved in determining whether the added benefits from this program justify the added costs. No short-term appraisal after one, two, three, or four years can be accepted as a really adequate indication of this. Most of the benefits of an education accrue both to individuals and to society over a period of years, and differences discerned at the close of the college career of students may not have a long-term reliability. When this issue of the long-term benefits of education is raised, it becomes easy to accept the contention of many faculty members that it is simply impossible to evaluate a program except on a long-term basis. Usually they quote examples by chapter and verse of individuals who 10, 15, and 20 years after college suddenly realize the enormous benefits conferred upon them by a college education, with specific reference to a particular course or professor.

At times in these days it would seem as though the primary goal of an educational program is to encourage people to continue their education by a continuing association with an educational institution. More and more programs are being set up for adults catering to professional developments and interests, recreational interests, and specific skills for various jobs and avocations. The idea seems to be inculcated in individuals that the only way to learn something is by enrolling in a formal course or series of lectures somewhere. The implications of independent study as a means of developing an individual who will continue his own education, which may, indeed, include such structured experiences, but which, to a large extent, will be done through his own reading, thinking, and interchange with others, are largely ignored. And, in this emphasis on the continuation of an education, the use of one's education may, to some extent, be overlooked. Surely we are not educating individuals simply to encourage them to continue educating themselves or to continue in an educational program; rather, we are educating them to a point where they can apply and utilize their education and then, as they find deficiencies in it for dealing with certain matters of concern to them, seek additional education in order to more adequately deal with the problems they face. Attention to the use of one's education forces evaluation into a future-oriented mode both to keep the individual and the professor alerted to the fact that it is what a person can do rather than what he has learned that is important, and also to keep before everyone concerned evidence as to whether the student really is progressing in regard to the ability to apply what he has learned. Unless an innovative program makes some impact in this regard, it may have some difficulty in justifying its existence.

The tendency to think of an undergraduate program as a unitary liberal education (which it seldom is, perhaps seldom has been) encourages faculties to think in terms of esoteric objectives which can be stated with great passion, which are quite unclear as goals, and simply useless as a basis for planning an education. Could education be broken up into a series of interrelated but really distinctive and certifiable competencies? If this could be done, various segments of a program could be thought of as units which contribute to particular competencies, and performance at the close of this segment of the program could be appraised so that the competencies would be verified. These, in turn, could be the basis for progressing to new levels, but they could equally be viewed, particularly for adults or for individuals who find it necessary to terminate their education or interrupt their education for a period, as useful competencies which might have continuing significance for work or further education. In some areas, the identification of such competencies is not very obvious, but a person who has finished a sequence of three or four courses in statistics, accounting, or had some grouping of courses in economics or chemistry should have developed certain identifiable skills, modes of thought, or effective ways of dealing with certain types of situations or problems. These could be appraised, validated, and reported to other professors or prospective employers, as assurance that this person has attained a significant and useful level of mastery of certain competencies.

In a sense, the question that is being raised is: What are some of the long-run outcomes of significance in an educational program? Some of these, as already implied, may have personal and social significance. For example, a few years ago a group of faculty members in general education science courses agreed that a major objective of these courses might be to encourage students to continue reading in the science area with new insights and appreciation resulting from their experiences in the course. A follow-up of students after taking this first course indicated that most of them did not continue reading science materials, even rejected the idea, and about the only ones who did, in most cases, were students who were converted to taking a science major. In one institution, however, a professor who had made ingenious use of the materials in his science course had had such impact on students that groups of students who had taken no science since their freshman year continued to meet in voluntary groups to read popular or semipopular science articles, discuss them, and occasionally call in their former professor to help them in understanding the meaning and the implications of some new scientific research. There was clearly an impact of major significance which had endured at least three years beyond the experience itself. This would suggest that, in some very real sense, the persons involved had a different attitude toward science, would be more likely to continue reading about scientific developments, and be more aware of the impact of science on our culture. Acceptance of such a long-term result as this as having value would almost certainly result in a major change in the patterns of most of our undergraduate course offerings and experiences. But even in innovative educational programs, faculty members find it difficult to accept or to plan in reference to goals which are not relatively immediate to the things which they themselves do with students. And since they are likely to modify their courses from time to time, evaluation, in the long haul or even on the short term, may be rejected as invalid because it applies to something that is no longer operative.

Rigor versus Reality

The point has been made that few people really expect a program to be dropped or greatly modified because of an evaluation study. Laudatory evaluation studies, of course, tend to reinforce what exists, but critical ones generally generate so many and diverse reactions that much energy is expended in attempting to demonstrate that the evaluation is invalid rather than restudying the program. In almost every evaluation, because of the problems cited, flaws can be pointed out which cause suspicion in the minds of others who have been ready to accept it and which may ultimately lead to the discard of the study as irrelevant or misleading.

Another reason why evaluation is less effective than might be anticipated is that it must, to some extent, concentrate on factual evidence, whereas actual decisions involve not only costs but values, and values, in turn, involve as much unrestrained emotion as rationality. Unless objective evidence of the benefits conferred by an innovative educational program relative to other existing ones is provided, so that resource allocations can be justified on a cost-benefit basis, values other than dollars are likely to be very prominent in any decision about change or elimination of a program.

Complete objectivity and valid judgment may seem to require evaluators who have autonomy and authority. And yet one may question whether there exist evaluators who are completely objective. There are few people who do not already have some bias about the desired nature of an educational program. Even having the objectivity, they may lack the experience and competency, and especially the sympathy, which are necessary in looking at, appraising, and rendering a judgment about something which is atypical. Hence, the use of an outside panel to review a program and render a report on it remains attractive to some persons, but the selection of such persons presents difficulties. If those operating the program have the selection, they seek individuals sympathetic to their efforts and those who are critical of a program may well wish to select individuals who are known to be dubious about such ventures. Then, too, there is a question of whether the evaluator himself may not have some kind of allegiance, an allegiance perhaps which is a part of his bias which makes it difficult for him to be absolutely fair. An evaluator picked by an administrator feels some allegiance to that person. An evaluator selected by a faculty vote or by the administrator of a program and paid directly by their order naturally feels some obligation to them. And, finally, those who have had experience with such panels know that it is sometimes very difficult indeed to get a panel of judges or consultants to render a negative verdict on a program after they have been exposed to the enthusiasm and wiles of those who operate it. And so we may suggest that there exist a few principles which obstruct a completely adequate evaluation: (1) methodological virtuosity will always be sacrificed to relevance and utility; (2) very few programs remain completely static long enough to be given a final evaluation; (3) an evaluation of a program dating back over several years is always regarded as irrelevant by those who were responsible for the first few years, probably because the program actually has changed, or at least those involved feel that they are now doing something different or are doing the same thing better; (4) a contention that the object of evaluation has changed is usually valid and, when evaluation has been a significant aspect of the program, at least some portion of the change may be attributed directly to the intervention of the evaluation activity; (5) an evaluation conducted in such manner as to have no effect on a program will be seen as irrelevant and will be ignored, and one which involves observation and interaction in depth with a program will change the program; (6) decisions and policies in educational programs result more from gradual, often unrelated, changes in practice rather than from a confrontation of issues and a resolution of these on the basis of principles and evidence.

The Purposes of Evaluation

In a sense, this final section takes us back to the beginning of the chapter. An evaluation of a program may be undertaken with the intent of reaching a decision about whether the program should be continued. It may be undertaken as a basis for bringing about some improvement in the program. It may be considered as an integral part of the program because of a belief that evaluation is an essential educational experience and also an educational goal. Those who are concerned with efficient allocation of resources, particularly when resources are severely limited, would very much like to carry on an evaluation of an innovative program, with an eye to reduction of costs or possibly elimination in order that the resources be made available to traditional operations in the institution. Those who are within the program and committed to it would like to view the evaluation as, at most, having implications for improvement of the program, and they would be happy indeed if the evaluation were so convincing as to justify additional resources or the expansion of activity. In some institutions, the concept of marginal benefits, of the allocation of additional resources plays a role. An institution already committed to several experimental programs for undergraduates may raise the question as to whether additional dollars assigned to this type of program would pay greater benefits than dollars assigned to support the traditional program. Undoubtedly, the question is a wise one, and it may be one to which evaluation can make some contribution, but the lack of agreement on the nature of the educational benefits conferred suggests that the action is more likely to be dictated by the personal values of individuals in authority or by political action within the faculty.

Evaluation is perhaps too often seen as something applied to a static program to determine its effectiveness. Evaluation is probably more significant if it is regarded as a means of facilitating the continuance of a dynamic and flexible educational experience. The observation of innovative programs suggests that many of these solidify into a highly structured program within three to six years, losing their innovative and experimental character while remaining dominated by a group of enthusiasts who are convinced that no changes are desirable. These enthusiasts may include students as well as faculty and administrators. Perhaps the most significant concept of evaluation, then, in an experimental program is one in which evaluation, as a thoughtful consideration of the significance of educational experiences from day to day in reference to broader goals determined in part by faculty and in part by students themselves, is accepted as an essential element of the educational program. In such a program, students are encouraged–and perhaps even forced–to be critical of themselves, of their accomplishments, and of the faculty and the program which is presumably responsible for them. Though faculty members will readily accept the idea that they are trying to develop students who can make judgments, they sometimes find it very difficult to live with students who are, in fact, making them. But if the faculty and administrators in an experimental program can view it as one which is dynamic and continually changing in which successive generations of students can continue to have the experiences which only happen to the first class or two in many innovative programs, then they may find that there is no conflict between teaching, learning, and evaluation but, rather, that evaluation is both the means to the improvement of teaching and learning and the end result of it.

Chapter 2

RESEARCH FOR ACTION[1]

Arthur W. Chickering

Until the late 1960s the typical college or university was complacent about its role, its performance, and its survival. Daily operations pushed questions of purpose and function into the shadows and there they stayed. Steady production of professionals and technicians blunted critics' probes; expert consultation and research services soothed anxieties about performance. But now students and parents, legislatures and foundations, and the colleges and universities themselves, are confused and uncertain about higher education. The quality of performance is sharply questioned. Some doubt that survival of present institutional forms is justified. The spotlights of public inquiry and popular pressures search the shadows and ask, "What are your purposes? What is your role? Why are such practices followed? What evidence justifies your existence?" These questions, and the mounting pressures behind them, must be answered with clearly stated objectives and pertinent evidence, the outcomes of sound evaluative research.

But evaluative research not only serves external interests. Students, faculty members, or administrators can disrupt an institution and can severely damage one another–physically, emotionally, spiritually. How will differences be resolved? Through power plays by polarized groups? By decisions

---

[1] Support for early portions of this manuscript came from the Project on Student Development in Small Colleges (NIMH Grant No. 14780-05) and from the Union for Experimenting Colleges and Universities. The final version had the benefit of critical comments and clerical assistance supplied by the Office of Research of the American Council on Education.

flowing from emotion in the service of prejudice, preference, and comfort? Or will reason, evidence, and objective analyses, well leavened by direct experiences, create conditions which foster individual development and institutional effectiveness? Currently, actions most often flow from emotion and uninformed opinion. To shift toward reason we need better ways to generate and to use pertinent research and theory, and we need more persons with the capacity to do so. More important, however, we need to learn better ways to think about higher education, its problems, and its potentials. For in the long run, *how* we think about problems is more important than *what* we think on a particular occasion.

John F. Kennedy said:

> The great enemy of truth is very often not the lie–deliberate, contrived, and dishonest–but the myth, persistent, persuasive and unrealistic. Too often we hold fast to the cliches of our forebears. We subject all facts to a prefabricated set of interpretations. We enjoy the comfort of opinion without the discomfort of thought. What is at stake is not some grand warfare of rival ideologies which will sweep the country with passion but the practical management of a modern economy. What we need is not labels and cliches but more basic discussion of the sophisticated and technical issues involved in keeping a great economy . . . moving ahead [Schlesinger, 1965].

Substitute "education" for "economy" and the statement still stands. In more stable and less complex times, when change was measured by generations, not years, the opinions of forebears reformulated in the light of personal experience could serve. But that process is no longer adequate. If American educational institutions are to be effective, they must forego the comfort of opinion and bear the discomfort of thought.

Evaluative research exemplifies one way to think about, and to come to grips with, educational planning and problem solving. But this approach has not yet been seriously applied in higher education. Except for Learned and Wood's (1938) early study of achievement and Newcomb's Bennington study (1943), most of the research on higher education has been reported during the past 10 or 12 years. The methods, the instruments, and the frames of reference which make sensible interpretation possible have only become available during the past four or five years. Moreover, because most researchers have kept the research process close to their vests, it has contributed little to educational planning and to daily decisions. Researchers have been more interested in theory testing than in tackling practical problems, they have cultivated and nurtured their own jargon, and they seldom have involved administrators, faculty members, or students–the potential consumer–in the process of planning, implementation, and analysis. As Sanford (1970) observes, practices which separate research from action not only lead to irrelevant research, but also cut the researcher off from investigation necessary to the development of his science.

The power of systematic evaluative research and careful application of results has yet to be demonstrated. A few short-term examples of limited scope can be cited, but sustained efforts backed by substantial institutional time, resources, money, and energy have not yet been reported. However, useful evaluative research, drawing on the growing pool of findings and experiences, is becoming possible. More research persons are interested in concrete applications, more are trying to use language understandable outside the family, and more are ready to work with students, faculty, and administration in the processes of evaluative research. Most importantly, many colleges have begun to take seriously problems of self-definition and self-evaluation.

Tomorrow's college, therefore, will walk on two legs–Action and Research. Sustained progress and reasonable equilibrium depends upon the complementary exercise of both. Hop along on one, stumble, and fall. That's the current pattern. Action ignoring, or ignorant of, pertinent information about students, institutions, social conditions, and their dynamic interactions can be little more than hit or miss. Even with the best information, each decision remains essentially artistic, resting heavily on insight and intuition. But insight and intuition starve unless nourished by new information.

The following pages suggest ways to generate needed information, ways which are particularly appropriate for institutions of small or moderate size, or for relatively discrete subunits of larger institutions. These "how-to-do-it" suggestions do not specify a required program to be undertaken all at once, but they do suggest some diverse possibilities within different areas from which selections can be made depending upon institutional priorities and resources. The suggestions fall in six general areas: clarifying objectives, assessing the environment, assessing student change, assessing attrition, priorities and scheduling, and appraising college impact.

The approach to evaluation implicit in the suggested alternatives rests on five major principles:

1. The information collected should have immediate, or at least reasonably contemporary, usefulness. It should not be necessary to wait two, four, or six years for the fruits of evaluation to be available and to have meaning.

2. The information should have high "face validity." It should be readily apparent that the information collected is pertinent to the purposes and practices of the institution and its students.

3. The information should be susceptible to simple, concrete, and easy management. Useful analyses and interpretations should not require complex statistical analyses, subtle clinical skills, or extensive training.

4. The information should be useful for both program evaluation and self-evaluation by students, faculty members, and administrators.

5. The information collected for immediate evaluative purposes should also be suitable for more complex longitudinal analyses.

These are challenging principles. The specific suggestions offered below do not always conform to them, although some come close. With further effort and experience, better approximations surely will be developed.

Clarifying Objectives

Some colleges and universities neglect any mention of institutional objectives or purposes; the simple fact of existence, together with descriptions of buildings, curricular requirements, general rules and regulations, and a list of faculty members and administrators which gives their major credentials, is judged sufficient. Other institutions nod briefly and go on to more pressing concerns: "The purpose of the College is to provide sound instruction in basic subjects, while encouraging each individual student to self-improvement and social usefulness. To this end, Ivy is using the income from its endowment funds and is enlisting the support of graduates and friends." Some colleges speak at greater length:

> **Hilltop College seeks to prepare men for lives of service, responsibility, creativity, and joy, both during and after college.**
>
> **The College shares with other liberal arts colleges of academic excellence–**
>
> - **a commitment to open inquiry by both its students and faculty, combined with rigorous appraisal and use of the results of that inquiry;**
>
> - **an emphasis on a broad education in the natural and social sciences, the humanities, and the arts, combined with strong competence in at least one field of the student's choosing;**

- **an educational program that aims more at preparing men to think and act clearly, boldly, and humanely in whatever life work they choose than at training for specific professional fields.**

The College's distinctive character comes from its striving for–

- **candor, simplicity, joy, and moral integrity in the whole of college life;**
- **a harmony for each man among his intellectual, physical, social, esthetic, and spiritual concerns;**
- **a creative use of smallness that places students in the closest contact with dedicated scholars;**
- **a sense of community marked by lasting concern of one person for another and by shared responsibilities for helping the College achieve its highest aims;**
- **a system of responsible self-government in the student body and in the faculty.**

Statements of objectives, when offered at all, are utopian and abstract. They mix hopes for institutional practice with objectives concerning student development; they aim to offer "sound instruction," "contact with dedicated scholars," "sense of community," "responsible self-government," and assume that these will foster "self-improvement and social usefulness," "harmony for each man among his intellectual, physical, social, esthetic, and spiritual concerns."

Clarifying utopian abstractions, and distinguishing between objectives for institutional development and for student development, not only sets a framework for evaluation, but also lays a cornerstone for educational improvement and financial survival. Institutional effectiveness improves when objectives are stated in concrete terms. Efforts to clarify objectives and to increase self-consciousness about them can release strong forces for institutional improvement and for student development.[2] The problem is to tackle the task and to complete it with some sense of success, with increased clarity instead of increased confusion. General discussion, good for opening the issues, almost universally frustrates attempts at satisfying closure or agreement. Following are some ways to make objectives more explicit, to move from discussion to a better basis for clarification, and to develop operational definitions of objectives which can provide a framework for evaluative judgments concerning institutional practices and student change.

---

[2] See A. W. Chickering, ***Education and Identity***, Chapter 8, "Clarity and Consistency of Objectives," for more detailed discussion.

Objectives for Student Development

An educational institution should continue to exist if it fosters aspects of student development which would not otherwise occur, or would occur significantly less elsewhere. Student change, therefore, and its relationship to institutional practices and conditions, is the primary criterion for institutional effectiveness and social contribution.

Because large numbers of students require postsecondary education, numerous institutions are needed, many of which will serve similar purposes. It is not necessary that each institution be unlike all the others. Each college need not be "distinctive" to survive, but clear stipulation of purposes and effective performance are required. A "distinctive" institution which hampers student development, or does not materially assist it, should not exist, even if sufficient financial support is available.

Three simple sets of activities can help clarify objectives for student development and set the basis for judgment concerning institutional effectiveness. I call them the ***Flesh and Blood Approach***, the ***Words and Deeds Approach***, and the ***Instrument Approach.*** There's nothing very new about any of them, and I cannot trace their lineage. I do know that *Methods in Personality Assessment* by Stern, Stein, and Bloom (1956) is one progenitor and is worth further attention.

The Flesh and Blood Approach. The Flesh and Blood Approach generates a list of valued characteristics for graduates. Students, faculty members, and administrators are asked to (a) name three–or four or five–students currently enrolled or recently graduated who best approximate the ideal graduate, and (b) to describe the two–or three or four–criteria used for selecting those students.

To maximize the involvement of administrators and faculty members, all should be asked to respond. Economic and logistical considerations probably suggest that, because of their large numbers, students should be sampled. Random samples should be stratified to reflect the proportions of men and women at each grade level, and also, perhaps, to reflect other significant subgroups whose ideal graduate might vary–major areas of study, commuters and residents, for example. The survey should be undertaken in late winter or spring when, presumably, most persons have become as well aquainted as they will during that year.

A committee of three to six faculty members, administrators, and students receives the reports from each respondent. Each committee member independently groups the characteristics–the criteria descriptions–under the minimum number of large generalizations acceptable to him. Then committee members share results, examine for compatability, and arrive at whatever consensus is possible. Their report is then shared with faculty, administration, and students to sharpen areas of consensus and to clarify dimensions of difference. Concurrently with, or subsequent to these operations, a second kind of analysis can be undertaken if institutional resources and interest permit.

This second analysis simply compares the students who have been nominated–or nominated several times if more stringent criteria seem appropriate–with their non-nominated peers. Admissions information, academic records, participation in extracurricular activities, test-retest studies of student change, attrition information, or other data judged pertinent and useful–all can be grist for the comparative mill.

Taken together, these two analyses move two tasks ahead. First, they identify major characteristics the college aims to foster among students and they support these generalizations with the wide array of specific characteristics mentioned by individual respondents. Thus, abstract objectives are made more concrete and explicit. Second, the comparative studies document differences between "ideal" students or their real life approximations, and their peers. These measures then can be used as more empirical bases for assessing student change and college impact.

The Words and Deeds Approach. The Words and Deeds Approach asks respondents to observe student behavior for a week–or two or three–and during this period to note words and deeds which exemplify behaviors characteristic of the ideal graduate. Here there is no attempt to associate the behaviors with a particular student, and this approach can be implemented anytime during the year. Again, it is most useful to involve all members of the faculty and administration in the task if possible. Students also should participate and sampling again may be required.

Analysis is similar to that first undertaken for the Flesh and Blood Approach. A representative committee receives the responses, gives some order to them, and shares them with the other members of the institution. The Words and Deeds Approach produces generalized statements of behaviors, accompanied by a wide array of specifics, which the colleges value and aim to increase in frequency and breadth.

Flesh and Blood, and Words and Deeds, as they are carried out by the faculty, administration, and students, usually heighten awareness of institutional objectives. They more clearly identify areas of consensus and of difference. The final results stand as useful guides for institutional practice and policy. Students who share the objectives can work toward them more effectively. Students who do not share the objectives, or whose values differ at significant points, can make more conscious choices concerning their own development.

The Instrument Approach. The Instrument Approach uses a questionnaire or personality inventory to generate a profile of desired student characteristics. Members of the faculty and administration complete the selected instrument or selected items therefrom, as they would have the ideal graduate respond. Use of the whole instrument permits calculation of scores and comparison with existing norms information. It also allows direct comparison with scores obtained from students at various points in time. Both represent substantial advantages over simple item analyses. But the time required to complete a full instrument may be more than faculty members and administrators will give. Under such conditions, items judged most pertinent to institutional objectives can be selected. Responses to these items are compared with students' responses to the same items, permitting comparisons between the "ideal student" and those currently enrolled. If test-retest data are available, student change can be examined to see whether responses move toward, or away from, the patterns generated by the faculty and administration.

It would be useful to have samples of students also respond in terms of their ideal. But the dangers of biasing student responses to the instrument when it is used for evaluative test-retest studies preclude use of student respondents in generating the "ideal" profile. The general fruits of the ideal generated by faculty and administration should be publicly available for general discussion, and some bias in later student responses may result. But that price must be paid, for there can be no equivocation about institutional objectives and values. Openness and honesty are ethical imperatives, and they also are the soundest basis for student cooperation and participation in efforts at institutional self-evaluation.

Objectives for Institutional Practice

Administrators and faculty members at every college want an educational environment–curricular patterns, teaching behaviors, intellectual climate, sense of community, student-faculty relationships, peer relationships, extracurricular activities–which serves their objectives for student development.

But seldom are the relationships between institutional practices and objectives for students made explicit, and rarely are those relationships seriously examined. General discussion of those relationships keeps them salient and helps identify areas of potential strength or weakness. But again, closure comes hard. Areas of agreement and disagreement remain hazy, and thus standards against which institutional change can be evaluated are not established. Often, therefore, descriptions of an institutional "ideal" in terms of explicit behaviors and practices, and concrete conditions of existence are needed.

Both the Flesh and Blood, and Words and Deeds approaches can be extended to this domain. Students, faculty members, and administrators can nominate, from among those current on campus, an "ideal"–teacher, residence hall supervisor, dean, class, extracurricular activity–and can describe the criteria used. This procedure can be powerful because it turns abstractions into known and concrete persons, practices, and conditions. But it is also dangerous. Invidious comparisons often are clearly implied and threaten some persons. Nominated persons may become precious, or become anxious and unduly self-conscious, and in either case their performance may suffer. These approaches can work, however, where there are high levels of trust, self-confidence, and openness. Unfortunately, although such conditions existed in the past, especially at small colleges, they are withering fast under the fire of external pressures and internal polarizations.

An application of the Instrument Approach can be less threatening. Its more limited power may be balanced by reducing resistance to the outcomes and by the more sharply defined criteria for evaluating institutional practices which are produced. This application of the Instrument Approach simply asks students, faculty members, and administrators to describe the institutional ideal by completing one or more–or selected parts of–questionnaires or inventories developed to assess the college environment and educational practices. Several instruments have been developed for these purposes: The *Institutional Self-Study Survey* (The American College Testing Program, 1969), the *Inventory of College Activities* (Astin, 1968), the *Experience of College Questionnaire* (Chickering, 1970c), the *Institutional Functioning Inventory* (Educational Testing Service, 1968), and the *College and University Environment Scales* (Pace, 1963). It is important to emphasize that the task is not to describe some abstraction of the ideal college or university, but to describe the ideals for one's own particular institution, taking clearly into account the students it serves, the interests, competencies, and values of the faculty members, and the resources available. For ideals must vary depending upon these institutional realities.

The scores and profiles generated by faculty, administrators, and students can be examined for congruence and incongruity. They can be compared with information available from these same instruments for other institutions, and with preexisting information concerning the institution in question if it exists. They also provide a benchmark for assessing institutional change as it may or may not occur when different kinds of students are admitted, when new programs are developed, or when institutional conditions change.

TABLE 1

Assessing the Environment

	General Climate	**Actual Behavior**
Groups	**Inventory of College Activities** *Institutional Self-Study Survey*	
	College and University Environment Scales Institutional Functioning Inventory	Experience of College Questionnaire
Individuals	Interviews and Questionnaires	Interviews, Questionnaires, Records

Assessing the Environment

Table 1 indicates four major components for effective environmental assessment, together with instruments and techniques appropriate to each. Two kinds of information are required: (a) information about the general climate—the values, behaviors, practices, and experiences which students *judge* to be widely shared; and (b) information about the actual behaviors, values, and practices which characterize the daily life of the institution. These two kinds of information can paint quite different pictures. Students can report that "many students get drunk on weekends," or "eighty percent smoked pot," when reports of actual behaviors indicate that only 10-20 percent actually did so; students can believe that "no one studies much around here" when reports of actual time spent studying indicate substantial investment.

These differences between general impressions and actual behavior result from two things. First, there is what Warren Bennis delightfully calls "the Pinocchio effect." Perceptions are distorted and distended by the student's frame of reference; whether a student finds a college challenging or dull, liberal or conservative, warm or cold, depends to some extent on his own school, community, and home background. Second, general perceptions may often reflect the influence and power of a visible, vocal, and active minority. Recent history documents this phenomenon on a national scale, where, with the help of mass media, general perceptions of college students have been pulled strongly toward the behaviors and attitudes of a small group. Because the pictures painted by these two different kinds of information often differ, both kinds are needed.

The general impressions, the "press" of the college culture as it is perceived, and the daily experiences and behaviors, both influence student development. One or another kind of information alone may not only mislead, but also may not identify important sources of college impact. The *Inventory of College Activities* and the *Institutional Self-Study Survey* recognize this dual need and ask students to report general impressions as well as their own activities, experiences, and achievements.

Such information is needed at two levels, for groups and for individuals. An institution must continually balance "actuarial" considerations and concern for individuals. Because developmentally powerful conditions and experiences may often create disequilibrium and stress, environments that are productive in an actuarial sense may also show high frequencies of upset and trauma for some individuals. Take, for example, a highly demanding and competitive environment where there are strong rewards for individual initiative, achievement, and integrity, and where failure means elimination. Under such conditions 95 percent of the graduates might develop high autonomy, competence, and integrity, but consequences for the other 5 percent might be suicide, demolished self-esteem, and unrealistically diminished aspirations. Is any change called for? Should demands and competition be moderated, risking less development for the 95 percent, in order to decrease the incidence of suicide and emotional upset. Every institution, consciously or unconsciously, has struck some balance. To discover where that balance lies and to identify the significant environmental forces at work, both group data and more detailed information for individuals are required.

The Group

Instruments used to assess students' general impressions, like the *College and University Environment Scales* and the *Institutional Functioning Inventory*, should be completed in the winter or spring after freshmen and transfers have become acquainted with the institution. Because perceptions of the college usually vary with grade level, proportionate representation from all four grades is necessary. There may be other subgroups for which accurate representation is important when samples are selected, although in most institutions the number of areas controlled will necessarily be small because sample sizes drop dramatically as additional criteria are added. Usually only a few controls can be introduced for any single study.

Instruments on which students report direct experiences and activities can be administered any time during the year. Toward the end of a semester, trimester, or quarter students have current and recent experiences to draw on. Reports may be more accurate then than directly after a vacation or at the beginning of a new period.

These instruments, like most other questionnaires and inventories, are best administered to all students at the same time, either in the same location or through some specified framework. Completing instruments as part of registration or as part of advising days has worked well in some colleges, for example. When this approach can be used, it is far superior to mailing or other more indirect distribution and collection. It controls the conditions for responding and thus (a) avoids having one student do the work for several of his friends, (b) makes it more difficult for a group of students to conspire to produce systematic effects, and (c) reduces distortions resulting from informal conversations, rumors, and bull sessions. It takes much less time administratively, and students–granting their general resistance to the task in the first place–prefer getting it over with on a stipulated occasion to the heckling and harassment that accompanies most mailing operations.

Responses to such instruments typically are analyzed simply by computing the frequencies and percents for all respondents, and for whatever subgroups are of interest–freshmen, sophomores, juniors, seniors; divisions or departments; residence hall units or special arrangements; commuters and residents.

The Individual

Obviously, questionnaires and inventories can be used on a one-shot or recurrent basis for studying individual patterns of behavior or experience. However, if the group assessments suggested above are undertaken, they are better supplemented by structured or open-ended interviews where students can respond more freely and flexibly, where more complex information can be obtained, and where interacting patterns can be described.

Two different interview designs can be used. Combining both is ideal because each yields different but complementary information.

The one-shot cross-sectional study selects several students at each grade level, and within each grade level chooses students who represent different characteristics at entrance and different patterns of experience and emphasis while enrolled. Once the sample has been selected, these students might be interviewed each day–at dinnertime perhaps–and asked to recapitulate their activities, experiences, and reactions during the past 24 hours. Within this framework each session might focus on a single area for special elaboration–student-faculty relationships, academic activities, study habits, extracurricular interests, relationships with friends and acquaintances, residence hall experiences. A week is probably the smallest period of time worth covering and longer periods or additional contacts at other times might be helpful if resources permit. These procedures give rich information about the institutions at one particular point in time, as it is experienced by varied kinds of students, in varied programs, with varied institutional careers.

The longitudinal study starts with freshmen, or freshmen and sophomores, and also selects a sample to represent diverse characteristics and orientations. These students are seen once, twice, or thrice a semester throughout their college career. The one-shot study above can provide a starting point for such an effort, which then simply continues contact with the freshman or sophomore samples. The longitudinal study does not yield immediately useful information, but it is the best way to get some sense of the total college experience as an individual student creates it for himself and is influenced by it. It suggests patterns of shifting relationships between the individual and the environment and provides a basis for judgments concerning more generally shared patterns suggested by the cross-sectional studies.

But the rate of institutional change plagues this kind of longitudinal study as it does all others. By the time four or five years have passed, students and institutions may have changed sufficiently to make any particulars irrelevant, although basic processes of response and adjustment, and basic patterns of reaction, tend to endure across shifting circumstances. Interview studies reveal such patterns and therefore may have lasting value.

When planning and implementing interview studies, keep the numbers small. Small! Very small!! Interviews are not for "counting." That's what questionnaires and inventories are for. Interviews instruct us concerning the varied dynamics, interrelationships, and interactions among college experiences and student development. Such information is not available from questionnaires and inventories except in the very gross terms provided by certain kinds of statistical manipulations. So use interviews for what they do best–elicit live and detailed information about how it is to live through, exist in, and cope with the varied institutional settings and pressures which characterize the college. Don't collect so much information that no one will ever have time to read it, much less digest it thoroughly. Better to have eight good interview studies that are thoughtfully examined, than to have 40 or 60 which are read once and then set aside until there is sufficient time to really mine all that beautiful ore–time which rarely comes.

Assessing Student Characteristics at Entrance

Student characteristics influence college impact two ways. First, a student's most important teachers are other students. Friends and reference groups filter the messages from the establishment and the student culture. They amplify or attenuate the forces of curricula, faculty members, or institutional rules and regulations. They can trump the best teacher's ace and stalemate the most thoughtful and agile dean. Thus, relationships with close friends and peer groups, or subcultures, are primary forces influencing student development in college.

Second, the impact of an experience depends upon the characteristics of the person who encounters it. Students differ in significant ways that affect their responses to a particular institution and the peers and faculty members they encounter there. Responses to different subject matters and tasks, different conditions of stress and release, different levels of satisfaction and frustration also are affected.[3]

---

[3] See *Education and Identity*, Chapters 13 and 14, for further details on these two propositions.

For these two reasons–because students are the primary educational agents for each other and because college impact varies with student characteristics–comprehensive information concerning the characteristics of entrants must be collected. Such information is required for sound educational planning and it lays the foundation for assessing student change and college impact.

Educational Testing Service and The American College Testing Program both market surveys for use with entering students and in follow-up studies. The surveys of the Cooperative Institutional Research Program of the American Council on Education also collect comprehensive information and share it with participating institutions. No one of these surveys will be completely satisfactory for an individual college. Because of the compromises required by instruments designed for large numbers of diverse institutions, the wording of some items will be off-target and some areas of special interest will be omitted or given only scant attention. But the economies of time and energy which accompany the use of such surveys more than compensate for the compromises required. Most important, the local institution can compare itself both with national norms and with particular subgroups of similar colleges. Therefore, first use one of these general surveys. Then if it seems necessary, supplement that information with a homemade instrument, and add whatever additional measures are selected for longitudinal studies.

Testing of entrants is best undertaken during the initial orientation period. This is especially important when measures of attitudes and values are involved because verbal responses can shift very quickly as students pick up the local culture. Some measures which may be used for advising and placement are necessarily completed by all students. Other measures used for longitudinal studies or to obtain general information concerning the distribution of varied background characteristics need not necessarily be completed by everyone. Various sampling alternatives can reduce the testing time required of students and cut costs for data processing and analysis. The sampling alternatives and procedures ultimately used will depend upon the number of entrants, the design for longitudinal studies judged most appropriate, and upon other local conditions.

Assessing Student Change

The cross-sectional and longitudinal interview studies provide major insights concerning the individual dynamics which underlie student development at the college. Thus, these studies not only help assess the environment but also are valuable supplements to test-retest studies of change using group measures.

Complex longitudinal studies which test the same students at different points in time, and which combine inventories, questionnaires, achievement tests, and interviews have been the method of choice for evaluating college impacts. Because student change is the major criterion, continuing study of that phenomenon is necessary. But when institutions are changing rapidly, in "image" and in actuality, and when the characteristics of the entrants also are changing, the contributions of complex longitudinal studies are diminished. The value of such studies begins to increase as an institution pursues them continuously for successive groups of students, because then patterns of change can be associated systematically with variation in the characteristics of the institution and its students. But such payoff probably requires at least 6, 8, or 10 years and there is not yet sufficient experience to really judge how valuable such efforts may be or how much interest they will hold for the faculty.

Note the distinction between usefulness to the individual college and other kinds of value. Complex longitudinal studies in different kinds of institutions are necessary for the continued development of knowledge and insights concerning higher education and its consequences for human development. Multiinstitutional research of this kind is particularly important. Moreover, it is appropriate for some institutions to contribute to this more general good. But expectations concerning immediate usefulness should be realistic and resources should be allotted accordingly.

The main reason a single institution should allot some time and effort to longitudinal research on student change lies in the important questions such studies keep alive. For they are the central questions administrators, faculty members, and students should continually ask. "What changes are occurring at this institution and among those of us who are part of it?" "What are the consequences for student development of current patterns of teaching and learning, living and loving, fleeing or fighting?" "What gains occur?" "What damages are inflicted?" In short, longitudinal studies–especially when accompanied by periodic interim reports and widespread involvement in planning, implementation, analysis, and interpretation–keep the institution focused on the main reasons for its existence. When this value is kept salient, the pertinence of information from other studies of goals, institutional characteristics, and student characteristics is more readily recognized.

Longitudinal studies can follow several different "designs" depending upon the complexity desired and resources available. Three publications discuss thoroughly various design alternatives for evaluative research: "Experimental and Quasi-Experimental Designs for Research on Teaching" by Campbell and

Stanley (1963), *Applications of Methods of Evaluation* by Hyman, Wright, and Hopkins (1962), and *Evaluative Research* by Suchman (1967). Read one or more of these and then decide which approach best suits local purposes and conditions. The following example is only one alternative. It is posited for an institution with an entering group of about 1,000 students, and aims to optimize the ratio of information to cost.

Randomly divide the entering students into four groups–A, B, C, D,–and administer instruments according to the following pattern.

TABLE 2

Design for Longitudinal Studies

Testing Times	**Four Randomly Selected Groups**				**Approximate Number Tested**[a]
	A	*B*	*C*	*D*	
At Entrance (during orientation or registration)	A	B			500
During Spring Semester, Freshman Year	A		C		500
During Spring Semester, Sophomore Year	A	B			350
During Spring Semester, Senior Year	A	B	C	D	500
Number of times each group is tested	4	3	2	1	

[a] Numbers estimated on basis of 50 percent attrition over four years.

This design controls for variation in the amount of testing and provides sufficient numbers during the first two years and after four years for study of subgroups. Numbers actually studied will necessarily vary depending upon the local attrition rate.

Use at least one instrument in the cognitive domain and one in the affective. Both kinds of development occur in, and are effected by, institutions of higher education. Furthermore, the cognitive and affective are interactive and to understand change in one requires some knowledge of change in the other. So don't undertake longitudinal studies of student development which neglect half the person.

Assessing human development between the ages of 16 and 26 is still in its conceptual and methodological infancy. The instruments currently available, though not infantile, are far from maturity. They all have substantial shortcomings. They are easily criticized and offer obvious targets for the sure laughs which nourish local wits–and half-wits. But the information they provide, the insights they suggest, and the questions they raise make valuable contributions. The alternative is relative ignorance, uninformed opinions, personal prejudice, folklore, and self-fulfilling prophecies. Judiciously used, the available tools are a worthwhile investment. The most frequently used instrument for the affective domain is the *Omnibus Personality Inventory* (Heist & Yonge, 1969). Its relatively widespread use during the past few years permits comparisons with a diverse range of institutions. Generate both scale scores and item frequencies and study changes for both scales and items. Often changes on particular items will be more instructive or thought provoking than the more abstract scale scores. Undertake study of changes for different subgroups where numbers permit and where decisions for action may flow from differential outcomes.

The most frequently used instrument for the cognitive domain has been *Graduate Record Examinations Area Tests* (Educational Testing Service) which assess knowledge and competence in the broad areas of natural science, social science, and humanities. Educational Testing Service and others have carried out fruitful test-retest studies using the Area Tests at entrance and again after two and four years. This instrument–which combines all three measures in one package, usually completed in one long sitting–is expensive to use. Its very general nature also precludes studies of relationships between college experiences and more particular aspects of intellectual competence and cognitive development. Educational Testing Service recently developed the *Survey of College Achievement* designed to assess achievement in English composition, humanities, mathematics, natural sciences, and "social sciences–history." Test-retest studies using these measures have not yet been reported, however, so their potential value for longitudinal research remains unconfirmed.

The range of intellectual competence and cognitive skills among entering students poses serious problems for the development of such instruments. At some colleges substantial numbers of students score near the top, and therefore the test cannot reflect development if it occurs. At other colleges the test is so difficult in language and style as well as in substance, for substantial numbers, that the meaning of higher scores on retesting is unclear.

Various editions of the ***Mental Measurement Yearbook***, edited by O. K. Buros, describe many other commercially available instruments accompanied by critical discussions of their strengths and weaknesses. Consult this reference for ideas concerning particular areas of assessment pertinent to local interests.

The most sophisticated design means nothing if the basic data are unreliable–if responses or observations differ unsystematically from one time to another. Unreliability usually comes from four major sources:

1. From subjects–responses or behavior under observation may vary because of fatigue, motivation, or shifting attitudes.

2. From situations–the conditions under which responses are made or observations are gathered may vary from time to time or from group to group and thus produce results which reflect changes in the situation more than changes in the respondents.

3. From the instruments–some questionnaire or test items may be ambiguous and be read differently by different students at different points in time. If the instrument is a human observer, his reliability may fluctuate with shifting attitudes, standards for judgment, or boredom.

4. From processing errors–inaccurate recording, scoring, coding, and copying all can increase error in the data and work against intelligible findings.

But no data collection is completely reliable. No matter how simple the measurement and how exact the ruler, there always is error. The task, therefore, is (a) to achieve the best reliability possible given the subjects, situations, instruments, and processing arrangements available, and (b) to recognize sources of unreliability when they have operated and to modify procedures, analyses, and interpretations accordingly, as best one can. Occasionally the unreliability may be so great that the only course is to throw out the findings. The wise choice is to do so, rather than to judge a program on a spurious basis.

As test-retest data become available, two general kinds of analyses should be undertaken. One set examines "net change," where all students who entered at a particular time are considered together to discover whether changes do occur and to learn something about their general direction and magnitude. The other set of analyses selects significant subgroups and examines change among them to discover whether changes vary for the different groups. The particular subgroups will depend upon local interests, priorities, and insights: "low risk" and "high risk," vocationally oriented and academically oriented, residents and commuters, liberals and conservatives, majors in humanities, social sciences, natural sciences.

Both kinds of analyses must be carried out because significant changes for subgroups may not be revealed in the general studies of net change, and because under some conditions such general studies may lead to inaccurate conclusions concerning student development and college impact. In several recent studies, subgroups of students within colleges have been found to change in opposite directions. In some cases, contrasting changes balanced each other so evenly that no net change appeared when subgroups were combined. For example, in Stockwell's (1967) commuter-resident study, net change data revealed no significant changes on a measure of intellectual disposition toward the natural science. Study of separate subgroups, however, revealed a statistically significant increase for residents and statistically significant decrease for commuters. Because the numbers of commuters and residents were about equal and because the magnitudes of change in contrary directions were similar, when they were combined the mean scores at early and later testing were identical for the total group.

In other cases, change in one group may outweigh contrary change in another so that statistically significant net change appears. But the findings are misleading because a subgroup showed statistically significant change in the opposite direction. Chickering (1970a), and Chickering and McCormick (1970), for example, found highly similar patterns of net change in several strikingly different small colleges and these were similar to those found in other longitudinal research on student development: increased liberalism, increased readiness to recognize emotions in thought and action, decreased concern for materialistic success. Study of subgroups within the diverse colleges, however, revealed groups of students who were changing in contrary direction on all these measures.

It is important, therefore, to study both net change and change for underlying subgroups. Even so, findings must be interpreted cautiously. Just as differences in change from college to college may result more from differences in the characteristics of the students than from differences in the college, differences in change for subgroups within a college also may reflect more the initial differences in their characteristics than the differences in their experience.

The important point is that student development and college impact result from reciprocal transactions between student characteristics and college experience. Studying subgroups as well as net change provides a much more solid basis for understanding these varied transactions.

Assessing Attrition

Studies of those who leave early usefully supplement studies of net change and subgroup change among those who stay. What are the characteristics of the early leavers? How do they view themselves and their futures? Where do they go and what do they do? How do they change after they leave? These questions have been pursued and the results indicate the value of such research.

Characteristics of Leavers and Stayers

Assessing attrition begins with comparative studies of leavers and stayers to identify similarities and differences among them. "Among" is not a grammatical slip because leavers who transfer and those who quit, go to work, and do not return, usually are quite different from transfers, and analyses should make this distinction. Therefore, compare three separate groups: stayers, transfers, and those who have not returned to college. A simple returnable postcard permits separation of the two groups of leavers. Then admissions information and instruments used at entrance and for retest studies can be examined to discover distinguishing characteristics. Because distinguishing characteristics may shift with grade level, separate analyses of freshman and sophomore leavers, and of junior leavers if numbers warrant, is worth the extra time and effort.

Follow-up Studies

Follow-up studies are frustrating. The ratio of effort to result often seems unfavorable. But such studies, especially when carried out periodically over a six- or eight-year period, have special value for institutional evaluation and program development.

Mailed questionnaires or telephone interviews seek three general kinds of information:

1. Why did they leave? What circumstances precipitated the decision to withdraw? Who was consulted while they decided? What reactions, advice, or counsel did they encounter?

2. Where have they been since they left? What have they been doing? What are their plans for the future?

3. How have they changed since leaving? How do changes among those who transfer and those who went to work, compare with changes among those who stayed through graduation?

All three kinds of information instruct us about the college and its impact. Though withdrawal may be precipitated by conditions internal to the students, leaving is often a response to the educational practices, cultural values, and conditions for existence which characterize the college. What were these? Was the college helpful as the question of leaving was considered? Research on withdrawal (Hannah, 1969) suggests that college faculty members and administrators only become involved late in the process when the decision to leave has become well crystallized. The experiences and changes which follow withdrawal also are instructive. A mailed questionnaire which recapitulates some of the items used for longitudinal studies enables comparative analyses of changes occurring among stayers, transfers, and students who leave and do not reenter college. These analyses provide valuable perspective on the contribution of the four-year college experience.

Therefore, follow-up studies are usually worth the effort. But if questionnaires are used, they must be short and simple, the appeal for cooperation must be honest and attractive, and two or three "reminder" mailings usually are required if returns are to exceed 50 percent.

Priorities and Scheduling

Few institutions can undertake all at once the full array of activities which have been suggested. Indeed it is unwise to try. Better to build a program gradually, with careful decisions concerning priorities and with widespread participation in such decisions. Then as feedback and use demonstrate which kinds of information are helpful and needed and which are not, a sound program can be developed. Practical considerations–staff time, funds for outside services, institutional talents and resources, prior commitments and the interests of administrators, faculty, and students–also must seriously influence the rate of growth and the timing for introducing new components. The schedule following, therefore, suggests a sequence of activities during a four-year period which may represent an appropriate balance for some institutions. Others may stretch the process longer, or may omit one or another component altogether. Few institutions are ready to move more rapidly. Write in your own dates; Year One might be 1971, '72, or '73.

Spring, Year One

Begin clarifying objectives for student development with either *Flesh and Blood* or *Words and Deeds,* and objectives for institutional development with the *Experience of College Questionnaire.*
Assess environment by administering *Experience of College Questionnaire* or similar instrument which calls for direct reports of experiences and activities.

Fall, Year One

Administer change instruments to entering students.
Undertake cross-sectional interview.

Spring, Year Two

Clarify institutional objectives by completing CUES.
Assess environment by administering CUES.
Conduct one or two follow-up interviews with selected participants in cross-sectional study.
Retest Year One entrants.

Fall, Year Two

Examine differences among stayers, transfers, and those who left to work, for September, Year One entrants.
Conduct follow-up interviews with selected participants in cross-sectional studies.

Spring, Year Three

Retest Year One entrants.
Conduct one or two follow-up interviews.

Fall, Year Three

Administer change instruments to entering students.
Examine differences among stayers, transfers, and those who left to work, for September, Year One entrants.
Conduct follow-up interviews.

Spring, Year Four

Follow-up interviews.

Fall, Year Four

Assess environment with *Experience of College Questionnaire.*
Follow-up interviews.

Spring, Year Five

Retest graduating seniors who entered in Year One, and retest Year Three entrants.
Follow-up interviews.

Appraising College Impact

Appraising college impact is no simple task. If change occurs, is it more or less than elsewhere, in or out of college? To what can it be attributed? To the general genetic and social forces at work in our culture? To particular characteristics of the students enrolled which dispose them toward special kinds of development? To the general atmosphere and climate of the college–or to particular identifiable programs, policies, and practices? To distinctive aspects of the local institution which generate special kinds of change?

Data alone don't answer such questions. Mere observations are insufficient. They must be given perspective by findings from other sources, and they must be enriched and enlarged by theoretical views, by a set of abstractions concerning student development and its relationships to college experiences and college characteristics. In a mature evaluation, as in a mature science, theoretical abstractions guide observations and observations correct, modify, or refine theory. Therefore, as data become available and as the tasks of interpretation begin, locate pertinent results from other sources and search out or develop relevant abstractions or theoretical formulations which can both enlarge and be modified by the data at hand.

Five books currently available are good leads to past research and offer theoretical views worth scrutiny for application to local circumstances: *The Educational and Vocational Development of College Students* by Astin and Panos, *Education and Identity* by Chickering (1969), *The Impact of College on Students* by Feldman and Newcomb (1969), *Growing Up in College* by Heath (1968), and that early magnum opus, *The American College*, edited by Nevitt Sanford (1962). These sources report general findings and present theoretical views based on research undertaken before 1969. Any two of these volumes will lead to more detailed reports pertinent to local concerns, and will suggest persons who can supply information about more recent work. The ERIC Clearinghouse on Higher Education, at the National Center for Higher Education, 1 Dupont Circle, Washington, D.C., also can be a major source of information concerning literature and research pertinent to local interests. With findings and concepts in hand from such sources as these, local results can be interpreted more soundly and reported more meaningfully.

But college impact has not been effectively appraised when written reports are distributed, even when these are set in the context of more general research and theory. Effective appraisal requires thorough examination, debate, and assimilation by the college community–by students, faculty members, and administrators, and by interested trustees and alumni. Commonly, substantial energy and resources are invested in conducting research and in preparing reports. Beyond that, however, research utilization is left to shift for itself. But the use of research does not just happen. It is one thing to get a report on to the desk, and into the file, of an appropriate consumer. It is quite another thing to get pertinent results and theory into his head and into his working knowledge.

Richard M. Nixon (1969) said, "But even when good and relevant research is achieved, its utilization is not automatic. Dissemination is not utilization, and much more than 'Brief Summaries,' announcements of results, and general research conferences are probably required to translate good research products into good program results. A whole new system of research delivery and application needs to be developed. . . . [p. 10]" He was right, and those comments apply to a single institution. Careful planning and systematic effort are required if research is to be used for action.

There is now a substantial body of experience and knowledge to draw from when designing programs for research utilization and planned change. The best single reference currently available is *Planning for Innovation through Dissemination and Utilization of Knowledge* by Ronald Havelock et al. (1969). Therefore, employ this work and the references it suggests to formulate plans for use so that integrated and mutually reinforcing efforts are undertaken. Without such efforts, evaluative research may fill four more filing cabinets with data and stick publication feathers in a professional bonnet, but it will contribute little to planning and action.

References

The American College Testing Program. *Institutional Self-Study Survey.* Iowa City, Iowa: Author, 1969.

Astin, A. W. *The college environment.* Washington, D.C.: American Council on Education, 1968.

Campbell, D. T., & Stanley, J. C. Experimental and quasi-experimental designs for research on teaching. In N. L. Gage (Ed.), *Handbook of research on teaching.* Chicago: Rand McNally & Company, 1963, 171-246.

Chickering, A. W. *Education and identity.* San Francisco: Jossey-Bass, 1969.

Chickering, A. W. Civil liberties and the experience of college. *Journal of Higher Education,* 1970, **41**, 599-606. (a)

Chickering, A. W. College impacts on cultural sophistication. Mimeographed. Plainfield, Vt.: Project on Student Development, Goddard College, 1970. (b)

Chickering, A. W. The Experience of College Questionnaire. Mimeographed. Plainfield, Vt.: Project on Student Development, Goddard College, 1970. (c)

Chickering, A. W., & McCormick, J. Personality development and the college experience. Mimeographed. Plainfield, Vt.: Project on Student Development, Goddard College, 1970.

Educational Testing Service. *College Student Questionnaires,* Part 1. Princeton, N.J.: Author, 1965.

Educational Testing Service. *Institutional Functioning Inventory.* Princeton, N.J.: Author, 1968.

Educational Testing Service. *Survey of College Achievement.* Princeton, N.J.: Author, 1969.

Feldman, K. A., & Newcomb, T. M. *The impact of college on students.* San Francisco: Jossey-Bass, 1969.

Hannah, W. Withdrawal from college. *Journal of College Student Personnel,* November 1969, 397-402.

Havelock, R. G. et al. *Planning for innovation through the dissemination and utilization of knowledge.* Ann Arbor, Mich.: Center for Research on the Utilization of Scientific Knowledge, University of Michigan, 1969.

Heath, D. H. *Growing up in college.* San Francisco: Jossey-Bass, 1968.

Heist, P., & Yonge, G. *The Omnibus Personality Inventory.* New York: The Psychological Corporation, 1969.

Hyman, H. H., Wright, C. R., & Hopkins, T. K. *Applications of methods of evaluation.* Berkeley: University of California Press, 1962.

Learned, W. S., & Wood, B. D. *The student and his knowledge: A report to the Carnegie Foundation on the results of the high school and college examinations of 1928, 1930, and 1932.* Bulletin No. 29. New York: The Carnegie Foundation for the Advancement of Teaching, 1938.

Newcomb, T. M. *Personality and social change.* New York: Holt, 1943.

Nixon, R. M. In G. J. Goldin, K. N. Margolin, & B. A. Stotsky, *The utilization of rehabilitation research: Concepts, principles, research,* Monograph No. 6. Boston, Mass.: New England Rehabilitation Research Institute, Northeastern University, February 1969.

Pace, C. R. *Technical manual: College and University Environment Scales.* Princeton, N.J.: Educational Testing Service, 1963.

Sanford, N. (Ed.) *The American college.* New York: John Wiley and Sons, 1962.

Sanford, N. Whatever happened to action research? *Journal of Social Issues,* 1970, **26**(4).

Schlesinger, A. M., Jr. *A thousand days.* Boston: Houghton Mifflin, Riverside Press, 1965.

Stern, G. G., Stein, M. J., & Bloom, B. S. *Methods in personality assessment.* Glencoe, Ill.: Free Press, 1956.

Stockwell, D. Differences in personality change of commuter and resident students. Mimeographed. Plainfield, Vt.: Project on Student Development, Goddard College, 1967.

Suchman, E. A. *Evaluative research.* New York: Russell Sage Foundation, 1967.

Chapter 3

JOHNSTON COLLEGE: AN EXPERIMENTING MODEL

Pressley C. McCoy

A cherished dream of many educators is the hope that some time during their career an opportunity might arise to create a new model for education in the liberal arts and sciences. Such a chance arose at the University of Redlands when the administration and the Board of Trustees decided during the mid-nineteen sixties to continue growing in size through the creation of cluster colleges. The first, they decided, would be an experimental institution of no more than 600 students. The only other specification was that the college should deal with the challenges of an urban society and with international problems as well.

When the information above was communicated to me by President Armacost in January 1968, along with an invitation to explore the possibilities, I gladly entered into conversation with him. Through a document which had already been drawn up by the University Board of Trustees, it was clear that Johnston College would have genuine freedom to experiment since the Board of Trustees had delegated to a separate Board of Overseers all responsibility for policy making in the academic area including admissions criteria, methods in teaching and learning, curriculum, evaluation, and graduation requirements.

The governance model created was also experimental in that it attempted to steer a mid-course between the decentralized arrangement of the Claremont Colleges and the more centralized design of the University of California at Santa Cruz. Rather than create separate administrative offices for admissions, business, development, registrar, etc., the University of Redlands chose to expand these existing offices to serve two colleges, each according to its own needs and goals. The administrative personnel concerned would work under the direction of the chief administrative officer of each college. In this arrangement, the theory is that greater economy will be realized than in the creation of separate offices for each college. One task of evaluation will be to determine the validity of that assumption.

Johnston College was blessed at the outset in the appointment of its first Chairman of the Board of Overseers, Dr. Dwayne Orton, who had gained distinction both in the educational profession, as a past President of Stockton College, and in the business community as Vice President of International Business Machines Corporation. In extended conversations with Dr. Orton, it became clear that Johnston College would afford its Chancellor the opportunity to create a new living-learning model which could attempt to be fully experimental. One severe handicap was the timing, for there was but one year to plan the academic program and appoint the faculty. A second limitation was financial, since the amount allocated for the planning budget was $100,000. Two dormitories and a dining hall had already been designed and were to be built with the help of federal loans.

The Academic Design

It was to the credit of the University of Redlands Board and Administration that they allowed the Chancellor complete freedom to design and plan the college program in his own way. The inadequate lead time was partially overcome through the invaluable experience provided by my opportunity to work with over 200 colleges in a 12-year span while serving as Director of the Danforth Foundation's Workshop in Liberal Education and later as President of the Central States College Association. Such exposure had stimulated thoughts about new ways of organizing a community of learning. Prior to the first consultation, I jotted down the following prerequisites:

1. Johnston College should focus its concern upon the growth of the person in the social context.

2. To keep attention directed to personal and societal needs, create an organizational pattern and choose those methods which would hopefully reinforce this tendency.

3. Do not establish departments, divisions, course units, nor grades which tend to glorify subject-matter fields as ends in themselves rather than as means to human ends.

4. Appoint faculty to the College but not to any department, division, nor dimension of the program. Allow them freedom to use the full range of their expertise and evaluate their role as facilitators of learning.

5. Establish as a condition for all faculty, students, and staff (including Chancellor) that each will accept full evaluation by all accepted methods and from all who have a valid basis for an opinion about one's work. Willingness to try Laboratories in Personal Growth under professional leadership will be a condition for all.

6. Primary emphasis on problem solving within the context of seminars (ratio of up to 14-1) and tutorials (ratio of up to 5-1).

7. Continuing and pervasive attention shall be given to the problems of human survival and higher quality in living. Examples would be genetic evolution of man, ecology, destructive conflict, scope and accelerated rate of change, personal and social integration, preservation and extension of basic human rights, etc.

8. Willingness to experiment on a continuing basis.

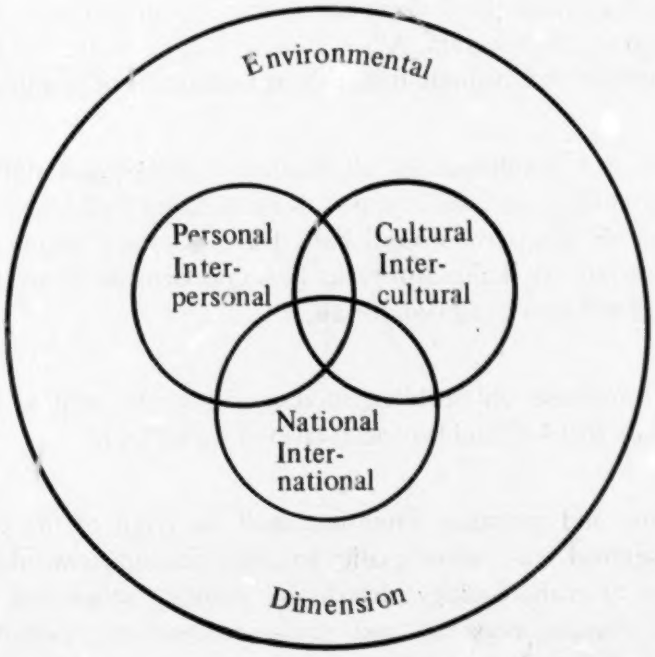
9. Viewing learning as a process involving the whole person with affective and cognitive aspects inextricably related.

10. Conscious concern for attitudinal change and growth.

11. Restore the old concept of co-learning which prevailed between master and apprentice or senior and junior co-learners.

12. The authority which counts is derived from genuine competency, wisdom, and concern. Authority by status or title may not be valid even though it has an impact on persons and processes.

13. To enhance the establishment of the preceding points, create dimensions of living and learning as points of reference and as a reminder that meaning inheres in the relatedness of persons to each other and with their environment. The dimensions which overlap and interpenetrate are as follows:



(It should be noted that the Environmental Dimension was added as an explicit and pervasive concern after faculty and students arrived on the scene, but much attention was given to ecology by consultants during the year of planning.)

Dimensions of Learning

A basic assumption of the Johnston program from the outset has been that meaning inheres in relationship: in encounter with oneself, other groups, nations, and the environment. Four dimensions of relationships–interpersonal, intercultural, international, and the dependence of all persons and groups on the environmental–provided the foundation on which the academic structure was first built. The intent has been to provide a problem-solving, creative, and practical organization to the student's course of development, leading from the most immediately personal matters to the largest concerns.

Interpersonal Dimension

The purpose of the Interpersonal Dimension is to help each student realize his full potential as a person, with a continuing concern for the determination of those values which form the basis for attitudes and decision making. In this dimension the academic focus is primarily, but not exclusively, on the fields of clinical and depth psychology, counseling, psychiatry, philosophy, theology, research on learning and creativity, discoveries in genetics and human ecology, and insights gleaned from literature and history. Seminars and tutorials in this phase of the program should help the student understand the dynamics of personality growth, interpersonal dynamics in one-to-one and group encounters, learning theories relevant to his own development, and constructive uses of stress, anxiety, and conflict. The theory and methods used in this aspect of the program draw heavily on the life work and publications of such noted scholars as Carl Rogers, Abraham Maslow, Viktor Frankl, Nevitt Sanford, Joseph Katz, Karl Menninger, and Erik Erikson. The insights of these authors are well appreciated by Johnston faculty.

Intercultural Dimension

The major goal of the Intercultural Dimension is to enhance our ability to live with the racial, ethnic, and cultural diversity that characterizes the world community. A major objective is to cultivate understanding of, and an ability to live effectively with, a culture different from one's own. Through this effort, the student will gain insights and competencies to improve his effectiveness within his own culture as well. Methods common to such academic disciplines as sociology, anthropology, and social psychology are integral to problem solving in this dimension.

One program within the Intercultural Dimension provides an integrated cluster of projects geared toward recruiting highly motivated minority youth; helping with their special educational, cultural, and counseling problems as they enter higher education; motivating and training them (through academic and practical field experience) for careers which serve their home or similar communities. The components of this program are:

1. College-community motivational programs. This aspect of the project is created to serve minority and/or disadvantaged communities (especially their own) to stimulate and support their aspirations for improvement of their communities and education for their youth.

2. College transition program. Offered here is a program of remedial and special seminars and tutorials, as well as counseling, to assist disadvantaged students in their transition to college life. The college transition program is also designed to encourage the contribution of these students to other intercultural programs of the college as they affect other students from both majority and minority backgrounds.

3. Bachelor's degree program with community development or social service emphasis. This phase of the program involves recruiting and training highly motivated minority and/or disadvantaged youth for professional and leadership roles in community work, and developing an undergraduate degree program to reach this goal. Career fields for which students will be prepared include teachers, intergroup relations specialists, school/home liaison workers, and in related fields such as social work, juvenile work, criminal justice, rehabilitation, and probation. All of this work must be accomplished within the broader context of the BA degree.

4. Student traineeships in community development. Trainees are paid apprentices in active community work and are drawn from minorities in the population. They work with existing agencies and organizations, and/or college-sponsored projects. Arrangements have already been made with some 30 agencies in southern California. Trainees will be paid for each year of their traineeship in lieu of scholarship grants.

5. Internship program. This part of the program is open to all Johnston students for academic credit as the internships relate to academic contracts. Arrangement is made for students interested in community development as either a career or a citizen's activity to affiliate with a public or private organization such as a school, business, or government agency for a semester or more.

6. Community insight. An intensive cross-cultural, family-based experience is available to all Johnston students in communities whose ethnic and social character is different from their own.

The projects described above will be carefully appraised to see whether the program stimulates and channels motivation of minority youth and others toward community development and/or social service careers. Notice will also be given to the degree of change in attitudes toward other cultures on the part of the entire college population. Johnston College is fortunate in having dedicated faculty and staff who, under the able leadership of Dr. Roger Baty, are assisting all students in dealing constructively with the intercultural dimensions of man's experience. We are also fortunate in being geographically located in a milieu of Black, Mexican-American, and American Indian culture groups. The intercultural phase of the program has been given a great boost through a major grant from the Danforth Foundation.

International Dimension

This dimension provides a cultural and linguistic framework for the student to develop his personal and academic interests, and focus them on critical problems of international relationships. The program is concerned not simply with the study of foreign cultures, but with establishing a global context in which one's own country is seen as a part. After examining his own cultural identifications, the student will develop academic skills to apply to the study of other nations. Seminars and tutorials in this dimension would draw on methods of such disciplines as history, law, political science, economics, psychology, etc. Since a certain percentage of Johnston faculty and students comes from nations outside the United States, problems of an international and intercultural nature are considered in a context of personal involvement. Field programs abroad will also help in this regard. For example, one faculty member will be doing a semester of research on the Alaskan Indians with the help of students, while another whose specialty is Chinese History plans to spend a year with some 12 students in Asia and the Middle East.

Environmental Dimension

In the Environmental Dimension, the student is helped to become aware of man's effects on his surroundings and his environment's effects on him, and urged to translate this awareness into responsible action. Students engage in problems of human ecology, environmental quality, nutrition, world resources, population, urban environment, computer usage, and creative artistic media. These interdisciplinary studies naturally involve a grounding in the fundamental fields of mathematics, physics, chemistry, biology, psychology, political science, sociology, and communications. Work in this dimension should help the student develop a sense of stewardship in the use of the earth's limited resources. Ideally, it should help him anticipate and cope with the impact of technological developments on the environment. Problems such as the conservation of energy supplies and aesthetics in use of our environment will likely persist for as long as man inhabits the planet.

With respect to the four dimensions outlined above, it is important to keep in mind that a given faculty member may offer seminars and tutorials in any or all of them during the course of a year. The only condition for involvement is his professional expertise and the expressed interest of students, as well as his own, in the problem defined. While a budget is created for programs which tend to fall within dimensions, funds are also budgeted for programs and projects which individual faculty wish to pursue not directly related to dimension programs. Such an arrangement helps prevent dimensions from assuming the less desirable characteristics of departments or divisions. All faculty have a hand in budget formulation and control. Directors of dimension programs serve on a rotating basis although exceptions can be made where faculty agree that program continuity and quality would suffer with change in leadership.

Learning Methods

We have made a deliberate attempt in the Johnston community to emphasize learning rather than teaching as the prime concern. Learning is a lifetime process for all persons who wish to continue growing including students, faculty, administrators, secretaries, maintenance personnel, and Board members. As we are able to develop evaluative procedures, we hope to appraise each person in the community in terms of his contribution to or facilitation of the learning of others as well as of himself. For example, the administrator who makes appropriate decisions and helps establish effective procedures helps facilitate a learning environment. Nothing deters learning more than the appointment of dull, ineffective faculty; and the administrator signing the contract is held accountable along with any others participating in the decision. To the degree, then, that learning objectives can be explicitly stated and evaluated, the task of assessing the relative contribution of those involved in the process becomes more realistic.

By keeping the focus on learning, a community need not engage in polarized debates about teaching versus research. It is obvious that a faculty member cannot learn if he does not continue to study and do research. Each faculty member is expected to organize at least one seminar or tutorial around his own research interests and invite students to work with him in the encounter. Some of the most exciting learning has taken place in these contexts. Learning is as much a necessity for the faculty member and administrator as for the student. A college which operates on this assumption automatically reduces the artificial polarization between faculty who teach and students who learn. In most instances the only distinction between faculty member and student should be one of relative depth and competency in the use of learning and problem-solving methods and in the amount of information stored in brain tissue. This, not age or title, is the source of the faculty member's true authority and the foundation for whatever academic respect is proffered.

The "Contract" Concept

Each semester, and prior to the January interim, students and faculty negotiate the curriculum to be offered on an individual basis. Students list the problems and topics they wish to pursue and the faculty post their own preferences as well. Boundaries for the negotiation are determined by the competencies of the faculty. Second- and third-year students who have approved graduation contracts already in hand are naturally in the position of persuading faculty to offer seminars, tutorials, or guided study in areas related to their personal graduation plans.

The contractual procedure helps the student learn to set his own goals and to be responsible for their realization. Rather then be told what he must take as a first-semester freshman, he is simply asked, "What is it you want to do? What are your current interests? Concerns? Problems? Challenges?" Begin where you are and work out from there. During the early stages of each seminar or tutorial, faculty and students discuss the nature of the contract or mutual commitment into which they are entering. Each contract in written form includes the following elements:

1. Background of interest and commitment to this problem or subject
2. Objectives
3. Resources on which to call (reading list, faculty, etc.)
4. Proportion of time allotted to seminar or tutorial
5. Procedures proposed to meet objectives (project, journal, paper, etc.)
6. Methods proposed to evaluate progress toward objectives
7. Relation of this course and/or problem to graduation contract
8. Group contract (responsibility related to the other members of the seminar or tutorial)
9. Nature of faculty involvement—commitment
10. Parties to the contract

Thus far it appears to us that the contract-making process encourages a degree of self-understanding and responsible participation in terms of personal goals.

Although more time consuming, it is less mechanical than the externally imposed units of graduation requirements which involve residence limitations, course distributions and sequences, grade point averages, and various numerical concepts and objectives. Following basic plans "grown from within" it is our hope that the student will be better prepared for future goal setting and self-criticism.

The term "contract" does not imply a legal procedure but implies long-range mutual commitment. As students mature, and plans change, contracts may be renegotiated (whether within an individual seminar or for graduation), but the change must be by mutual consent of all parties involved. This method also reinforces the notion of personal accountability to other persons and groups.

If, during his second year, the student decides to pursue a Bachelor of Arts and Science Degree, he must ask his adviser to help him draw up a graduation contract which will then be presented by the two of them to a graduation committee consisting of at least three faculty members and possibly one student. When the student feels he has fulfilled his graduation contract, whether in three, four, or five years, he requests to appear before the committee for examination. The graduation committee evaluates contracts and their fulfillment with the following criteria (drawn up by unanimous vote of faculty and students) in mind:

1. Personal integrity. Does the plan facilitate integration of parts of the student's experience so as to make or keep him a whole person? How does the proposal combine the cognitive (intellectual) and affective (behavioral) aspects of learning?

2. Integration of knowledge and mastery of varied methodologies. Does the plan include mastery of several important methodologies, or ways various disciplines approach the study and solution of problems? Are there means in the proposal for combining the knowledge of these and bringing it to bear upon a crucial problem or problems? Can the student show not simply that he "knows" some economics, or physics, or literature, but that he has "learned how to learn" about each of these disciplines–that he has truly mastered their methods and not simply their content?

3. Language. If a foreign language is not included in the proposed contract, is there good reason for its omission and is there a reasonable substitute?

4. Contemporary problems. Does the plan indicate continuing awareness of urgent problems of our time, or does it ignore these problems? (These problems can, of course, be spelled out in detail, and probably should be.) Does it reflect concern for problems in interpersonal, intercultural, international, and environmental areas, as well as the way in which science and technology affect these?

5. Independence. Is there evidence that the student is increasing his ability to work effectively on his own?

6. Professional objectives. Does the plan take into account the student's professional or graduate school objectives?

7. Postgraduate education. Does the plan suggest a program of "lifetime learning"?

8. State requirements. Does the plan meet state requirements for special courses and for careers?

9. Physical education. Does the plan include competence in at least two sports that will have sustained value in the student's future?

10. Concentration. Does the plan have a feasible focal point, a "major" emphasis or its equivalent?

11. Cross-cultural experience. Does the plan include in-depth experience in living, learning, and working in a culture other than his own–either in the U.S. or abroad?

The contract plan for seminars, tutorials, and graduation attempts to provide the following assets: maximum flexibility in education; reliance upon student initiative and responsibility in creating and carrying out their own programs; guidance and control of the procedure by faculty members of diverse disciplines; and the removal of the traditional credit compilation and four-year residence requirements, while guaranteeing breadth in academic experience.

Laboratories in Personal and Social Growth

In developing this aspect of the program, we have drawn most heavily upon the work of the National Training Laboratories in Bethel, Maine, the Center for Creative Living and Spiritual Growth, and the Organizational Development Program of TRW Systems, Redondo Beach, California. Professional consultants have been drawn from these and other institutions under the leadership of Mr. Del Poling and Dr. Glenn Whitlock. In addition to the "encounter" laboratories, extensive use is made of psychodrama under the direction of Dr. Donell Miller. Every precaution has been taken to ensure that this training is managed with scrupulous care under the highest professional standards.

The personal growth labs and psychodrama have been incorporated from the initiation of the program because we consider human feelings, emotions, and imagination inseparable from cognitive processes and the performance of all human tasks. When a student becomes bored in a seminar, learning progress has stopped. Unless he feels free to express his feeling, much time and effort may be needlessly squandered. Hostility, anger, resentment may also be roadblocks to individual and group learning, quite apart from the other negative effects. Since the College avows a holistic philosophy and since we are reaching for maximum effectiveness in learning, we can hardly ignore the role which emotions play in human interaction. The laboratories are assisting us in establishing a climate of greater trust and clarity in communication. Students are feeding back in a constructive manner frank criticisms of faculty and administrators in the living-learning context. Once students learn that expression of negative feelings is not socially punished, they no longer need to resort to "masks" or withdrawal. Within the training group, an individual is invited–sometimes challenged–to quit playing roles and wearing masks in a defensive or destructive manner. His uniqueness is respected for what it is. As the individual makes himself known, and comes to know others, he increases not only his own sense of worth and self-understanding, but his basis of personal trust. He is able to affirm others and is confirmed by them.

Such a capacity for interpersonal relationships is valuable in all aspects of the academic program: in seminar involvement, in community process, in the field, and in planning individual projects. One purpose of the personal growth methods is to prepare the individual "for life now." The Johnston catalog sets forth the following conditions for the student to consider if he is to find improvement through change in insights, understanding, sensitivities, and skills:

> Presentation of self: Until the individual has an opportunity to reveal the way he sees things and does things, he has little basis for improvement and change.
>
> Feedback: Individuals do not learn from their experience. They learn from bringing out the essential patterns of purposes, motives, and behavior in a situation where they can receive back clear and accurate information about the relevancy and effectiveness of their behavior. They need a feedback system which continuously operates so that they can change and correct what is inappropriate.

Atmosphere: An atmosphere of trust and nondefensiveness is necessary for people both to be willing to expose their behavior and purposes and to accept feedback.

Cognitive map: Knowledge from research, theory, and experience is needed and important to enable the individual both to understand his experiences and to generalize from them. But generally information is most effective when it follows experience and feedback.

Experimentation: Unless there is opportunity to try out new patterns of thought and behavior, they never become a part of the individual. Without experimental efforts relevant change is difficult to make.

Practice: Equally important is the need to practice new approaches so that the individual gains security in being different.

Application: Unless learning and change can be applied to "backhome" situations, they are not likely to be effective or lasting. Attention needs to be given to helping individuals plan application.

Relearning how to learn: Because much of our academic experience has led us to believe that we learn out of listening to authorities, there is frequently need to learn how to learn from presentation-feedback-experimentation.

The initial laboratory experience provides opportunities to explore traditional patterns of behavior.

In the personal growth lab a vacuum is created by the trainer's refusing to carry out the traditional expectations of his role: leadership, agenda, and procedure setting. Into this vacuum of lack of formal leadership and lack of clarity about goals and procedures, members rush in with the purpose of filling in the missing ingredients. Thus the first condition of training is met.

Each laboratory has as its task improving the learning of each individual. For feedback to follow presentation of self, an appropriate climate needs to be developed. Legitimate opportunities for individuals to try out new ways of behaving need to be present. The training laboratory is designed to meet these conditions.

Evaluation Processes

The comprehensive, personal critique replaces traditional letter or number grading and is directly related to the seminar or tutorial contract. As the seminar draws to a close, both student and faculty member write a critique of the student's work with the learning objectives of the contract providing the points of reference. The evaluating process provides a learning opportunity for both student and faculty member in a manner not possible with conventional grading. As those who have tried such methods know, this process requires more time and energy than the quicker labeling process.

Judgments about the performance of students should not be oversimplified, arbitrary, nor dehumanized. Therefore, critiques are a matter for discussion before the final copy is placed in the student's permanent file. An effective critique discusses the range of strengths and weaknesses of a given work, enabling the student to see where added concentration of effort is needed. This approach also helps the student develop his own powers of self-analysis and criticism which must carry him through the future. The evaluation form covers the following points:

A. The following have been agreed to by faculty for content of evaluations:

  1. Name of student, etc.
  2. Seminar or tutorial description: objectives, both quantity and quality
  3. Learning performance (faculty and students), both cognitive and affective
  4. Deadline for evaluations will coincide with University College grade deadline

B. The following are additional desiderata:

  1. Unit equivalent of course (i.e., hours, etc.)
  2. Book lists
  3. Meeting times and places
  4. Extra lectures, films, etc.
  5. Attendance
  6. Completion of goals in contract
  7. Areas of competence (academic and otherwise) as a result of the seminar or tutorial
  8. Personal participation: attitude, initiative, personal involvement, reading, other
  9. Evidence of progress in understanding, insight, participation, independence, maturity, creativity
  10. Level of work as a consequence of seminar or tutorial
  11. Possible areas of further study as a result of course
  12. Strengths and weaknesses; recommended improvements

Should a student wish to transfer, "translations" of critiques will be made upon the request of the student and in consultation with him. This will be done, however, only if the college to which he is transferring demands it.

Through the contract and critique system, every student is treated as a potential honors student; each individual, whatever his position, is valued according to his expressed needs and evident potential rather than according to depersonalized standards of what the student or faculty member must be like. It is our belief that this system can be used with the university-wide ratio of 14 to 1. The present ratio at Johnston College is 12 to 1.

Field Work and Internships

All students are encouraged to include off-campus field work in their plans; field work can be either an individual or group project, related to any dimension of the program. It may occur regularly throughout a semester or year, or occupy an intensive short period of time such as the January Interim Semester. An example of the latter is the Community Insight Project explained previously in this chapter.

Internships of a semester or a year in agencies or institutions related to the student's concentration can become a part of the graduation contract, and might affect the duration of the student's enrollment in the college. One student, for example, is planning on a year of supervised work in a Japanese electronics firm. Careful preparation is made for both field work and internship experiences and both involve coordination between a faculty member and the person supervising the student in the work environment.

Evaluation of the Johnston College Program

It is my belief that every educational enterprise, whether it calls itself experimental or not, should do its very best to appraise all aspects of its program on a continuing basis. Evaluation is an integral aspect of the learning process. Unless individuals and institutions examine results in the light of avowed objectives, how can they possibly know wherein they are succeeding or failing? The tough part of the challenge comes in finding or creating adequate means of assessment for various objectives and making proper use of them.

All three of the purposes stated in the last portion of Dr. Dressel's introductory chapter apply to the Johnston context. The University Board of Trustees declared at the outset that the total program, including the governance arrangement between the College and the University, would be evaluated following the third year of operation. A five-year testing period would have been preferred since this would allow time for appraising the work of graduates at least one year beyond commencement. Not only is evaluation being used to determine whether the program should be continued, but we have already made some changes as a result of feedback. For example, during the first semester of the college's history, the faculty and students planned a Quest for Meaning (QFM) seminar which would involve the total community. Although the ultimate in flexibility was suggested among and within various seminars, some succeeded admirably while many seemed to flounder. When the causes were examined, it became apparent that varied abilities of faculty, combined with motivational variations of students, were the determining factors. This year a few of the QFMs are being offered and are succeeding. Had this seminar been conceived as an indispensable part of the design for a fixed number of years, learning opportunities would have been dissipated.

An evaluation consciousness pervades the Johnston community due to the contract-evaluation system for semester work as well as for graduation. Faculty enter the community well aware of the importance attached to this process, through the interview and the explicit statement in their contracts. During the first semester, all seminars and tutorials were evaluated by students, and the data collected formed the basis for recommending reappointment of faculty to the Board of Overseers. Since virtually all faculty were on initial one-year appointments, and since the data indicated a considerable measure of success in realizing learning objectives, all were reappointed. However, the healthy habit of expecting feedback was established from the beginning.

Baseline Data

Since the establishment of Johnston College represented a new approach to learning alongside a larger, well-established liberal arts college, it was important to use whatever instruments were at hand to establish baseline data for the purpose of later comparison within the University as well as with national norms. However, since the goals of Johnston are distinctive, the key question is to what extent these tests will tell us what we most need to know.

The Institutional Research Office at the University of Redlands administered to incoming students at both University College and Johnston College the Omnibus Personality Index, the College and University Environment Scales, the Whittaker Scale of Values, and an instrument developed by the American Council on Education for the attitude perception of students. In addition, two Johnston faculty in psychology administered certain personality tests and diagnostic, therapeutic instruments to ascertain the individual levels of mental maturity. Instruments used by them are the Edwards Personality Scales, the Minnesota Multiphasic Inventory, the Meyers-Briggs Personality

Scales, and the College Student Characteristics Index. Great care has been taken to guarantee that the confidentiality of all test results is not violated.

Evaluation by Visiting Experts

Since all instruments have their limitation, it seemed wise to arrange for continuing evaluation of the Johnston program by a team of visiting experts who would make periodic visits to the campus and conduct in depth interviews of students, faculty, and administrators. Dr. Nevitt Sanford, Director of the Wright Institute, and his colleagues have agreed to conduct the evaluation for at least a two-year period. The team of interviewers, after completing an initial period of information gathering, will focus upon three elements of the Johnston program, namely, the affective, cognitive, and intercultural. They state that their "evaluation will be based upon the objectives obtained through interviews with those members of the Johnston College administration, faculty, and staff who have particular responsibility for the development and implementation of these programs."

As a part of the evaluation and consultation assignment, the Wright Institute will interview selected samples of students from both Johnston and University Colleges whose OPI profiles closely approximate each other. Through this approach they hope to gain perspective on the relative impact of the Johnston program upon its students.

Representatives of the Institute will rely upon group meetings with faculty and administration as the primary means of reporting the results of their interviews and observations. Such meetings have already provided opportunity for discussion of findings by the persons most involved. The reports are scheduled for the beginning and end of each semester. The following time schedule reveals the extent and nature of the involvement of the Institute:

Evaluation Time Schedule

October 1970

Selection of nine pairs of Johnston College and University College students, each pair sharing a similar OPI profile, but with different profiles represented among the pairs.

A discussion/feedback session with the Johnston College faculty, administration, and staff regarding problems that have appeared at the outset of the second year, and improvements they have noted in difficult areas over last year. We will share in some more detail results of the interviews conducted with students and faculty at Johnston College last year.

Supplemental interviews will be scheduled with members of the Johnston College staff regarding the conceptualization, objectives, and integration of the affective program at Johnston College.

Interviews will be conducted with members of the Johnston College staff regarding the conceptualization, objectives, and integration of the Intercultural Dimension of the program.

Interviews with the matched pairs of Johnston College-University College students.

December 1970

Report/discussion of the affective program: Presentation by Dr. Sanford of his conceptualization of the terms and dynamics of adolescent development, with discussion of the problems of affective assessment, evaluation, and integration with the cognitive program.

Interviews regarding the contract-making process at Johnston with a sample of students (small) and faculty at Johnston College.

February 1971 (prior to making new contracts)

1. Interview matched pairs of Johnston College-University College students.
2. Interview students newly returned from their interterm intercultural projects, perhaps sitting in on the students' reports to faculty members.
3. Report/discussion on contract making and evaluation.

May 1971

Report/discussion on the interterm intercultural program.

1. Interview sample of Johnston College faculty, both first and second year.
2. Interview the student matched pairs.
3. Consult with Pressley McCoy and Edward Williams on their assessment of the year and projections for the next year.

October 1971

Interview student matched pairs.

Interview sample of students from first, second, and third year's classes.

At their request, consult with the Johnston College faculty, administration, and staff through the remainder of the contract period.

February 1972

Interview student matched pairs.

Final interviews with a sample of Johnston College students, faculty, and administrators.

May 1972 (Final Report of the Wright Institute)

Evaluation of the study of matched pairs of students.

Written evaluation and report on the Johnston College program.

It is hoped that the OPI, together with the interviews of the Wright Institute, will enable us to determine whether the Johnston program is any more successful than the conventional college in enhancing openness, risk-taking abilities, and other personal qualities conducive to learning and positive social action.

In addition to the above, the Institutional Research Office of the University is planning on administering the Institutional Functioning Inventory, developed by Educational Testing Service, to determine the faculty awareness levels of functioning ability. The Berkeley-Wilson Faculty Questionnaire, involving an interpretation of faculty perceptions of themselves as community members, will also be used.

The Omnibus Personality Inventory was chosen because it measures many of the most important personality dimensions related to college students. We believe it is very well constructed and validated when compared to similar tests. Its wide use also provides a comprehensive basis of comparison with a national norm as well as with individual colleges and universities throughout the nation.

Similar reasons apply to the College Student Questionnaire. It provides vital demographic information on the students and their background as well as some essential attitudes with respect to motivation for grades, liberalism, social conscience, cultural sophistication, family and peer independence.

In addition, we have available two academic or intellectual measurements which are the high school grade point average and the students' verbal and quantitative scores on the Scholastic Aptitude Test.

Mr. Kouji Nakata of UCLA, who has devoted one-half of his professional time to evaluating the Johnston program since its inception, raises the following fundamental questions regarding the research design:

> **The research program that was laid out was by most standards a sound and comprehensive one. Its faults, I think, are due primarily to the underlying assumption that what is being measured represents the appropriate indices of what defines a good education at Johnston College. Basically the issue is one of clearly defining the specific educational goals for students attending Johnston College and then designing a research program that accurately and fairly measures their achievement along these goals.**
>
> **After being involved with Johnston College for over a year, it is apparent to me that these traditional research methodologies are inadequate and misleading. In conversation with Dr. Paul Heist, one of the co-authors of the OPI, he also confirmed the fact that advances in higher education, as Johnston College represents, have moved beyond the research methodology needed to measure their effectiveness. This is not to say that such instruments as the OPI and CSQ are not to be used, but that they are not enough nor do they focus on the areas of primary importance. In this same position are the usual measures of academic achievement such as grade point averages and standardized test scores.**

Rapid changes in society have necessitated significant changes in educational processes. It may well be that adequate measures of new developments in schools and colleges will have to await the creation of instruments or methods not now available. Meanwhile, we shall do the best we can with what the practitioners of the art have given us.

To be successful, the experimenting college must have not only an evaluative minded faculty, but one which experiences personal and professional growth in the learning process. The following representative reports of the faculty on the first year reveal a sense of dedication to learning both for students and themselves. Their comments on strengths and weaknesses in various phases of the program may be of interest to the reader and are included with the writers' permission as submitted:

> **Dear Pres:**
>
> **I have learned more about my profession during this year at Johnston than during any other two years of my admittedly brief career. I'm very glad that I came to Johnston**

College with at least several years' experience, however; I think I would have been completely at sea had I not had some resources to draw on. This is especially true since I am a rather structured sort, and operate best when I have a fairly definite plan before me. My experience enabled me to propose some different models of course organization to my classes; most of these turned out to work very well. At ___ I was a lecturer by and large; "discussion" classes usually focused on student responses to my views, not ideas generated by the group. Four years of practice and an enormous ego combined to make me quite successful with this approach. Yet as I refined it, and as I taught the same courses again and again, I began to sense that my "practice" didn't really count for that much any more and that my "enormous ego" was taking over. All my teaching became more faculty-centered and less student-centered. I would change books in a course not so much for the benefit of students as for my own education. Rationalizations such as "I won't go stale if I teach Howells instead of London this year" prevented me from seeing that I was really searching for a new vehicle to display my gifts as a lecturer. I had adopted an adage from one of my colleagues—"It's not the man but the material"—but I was discovering that I really didn't believe that. Instead of trying to make dead books live for a student, I was largely trying to make "the man," myself, very much alive for a captive but admiring class.

Now at Johnston College these purposes are naturally still in operation; neither I nor very many of my colleagues are really selfless people, and I would never take a job which didn't give me some ego satisfactions. But I don't feel dominated by them as I did at ___. And the reason for this is that I'm able to do *more teaching and less leading*, I'm learning the pleasures of getting out of a student's way, of making myself useless, of watching growth instead of insisting upon it. In fact, the one persistent critical comment I received on my otherwise very favorable evaluations was that I did not talk enough, or that I was too self-conscious about "lecturing." This was particularly true the first semester; I was falling all over myself trying not to lecture anyone about anything. I'm now at the point where I can begin to appreciate a number of teaching "methods" (horrible word—sounds like a sex manual). A good teacher in the Johnston style has to develop a good ear, it seems to me. That ear will tell him when it is appropriate to lecture, when to shut up, when to ask questions and when to answer them. One of the beauties of the college is that it does not prescribe or even inadvertantly encourage one method of teaching. And it's here that I can return to the importance of experience. Even in a straight lecturing situation a teacher must be aware of when he loses the class, and when he "has them." And it takes time to acquire this gift; I sense that I am just beginning to develop mine. Good teaching demands seasoning, exposure, and confidence; these things are hard to come by in the first year.

Where I have been much less successful (fortunately, in some ways it's less important that I have such talent) is in mastering some administrative skills. Part of this is inbred prejudice. I grew up in a family of teachers, and when my father was literally seduced into taking a high school principalship for six years I found myself living with a man who heartily disliked every aspect of that job and who longed for a return to the classroom. My father is anything but a complainer, but he crabbed and snorted his way through those six years as if they were an undeserved penance for some hideous crime he could never have committed. So I've always looked on administrators through heavily tinted glasses; the bad ones are sadists and the good ones are martyrs. Accordingly, I try manfully to help in whatever ways I can with the incredible burdens that martyrs like you and Ed face, but I still find it very hard. I have more skills in those "traditional" concerns of faculty–graduation requirements, contracts, hiring and firing, and the like. Despite my background in philosophy and theology, I've never had much interest in a general philosophy of education. I admire the way that you can pull out an answer to the inquiries of the Board; it would take me an hour to formulate my views on similar problems. I'm much more interested in *doing* education–that's where my talents lie.

I'm only beginning to appreciate my relationships with students in the Johnston College atmosphere. I imagined (with grounds, I think) that I was one of the most open and least dictatorial teachers at___, and that this was one reason for my many friendships with students. But from this perspective I can see how many of my aggressions I really was expressing through my teaching. I think rather guiltily of some of the people I failed, of some of the exams I gave scant reading to; of some of the caustic comments I embroidered on papers; the more students you have the easier it is to think of them simply as "students." I suspect that some of my student friendships at____will be among the strongest of my life simply because they were involved in my first four years of teaching; I entered and graduated with the class of 1969. But the framework of openness and verbal as well as written confrontation (the latter can't help but be easier) which is built into the Johnston College program offers the greater potential for lifetime friendships as well as lifetime learning.

So I'm afraid that despite all our good intentions we can't help but be a bit clannish as a faculty next year–perhaps even smug. It seems to me that we've accomplished a great deal for the first year of a college, and that we're inevitably going to have a slightly possessive feeling about it. I hope I've turned the "ego corner" in teaching, that our experimental thematic approach to an English major (I'll run two seminars in the fall) will be as successful as students seem to think it will be, and that I can continue to grow in tuning my ear to the stated and unstated needs of students.

Fellow in English Literature

Dear Pres:

I would like to make some observations and evaluations from the past year and some suggestions for the next year. Some observations are undoubtedly repetitive, but it may serve some purpose to put them down together. The suggestions for the future in the Interpersonal area will be considered in programs planned for the fall and for the remainder of the year. In general, although I have had disappointments both in the college and in myself this year, I am strongly optimistic about the future. Perhaps we should expect to make more than the average number of mistakes for an academic institution, but hopefully we will learn more precisely because our failures are more gross. Likewise, our successes seem to be more spectacular. And if Nevitt Sanford is correct, that a college gains stature by its successes rather than by its average productions, then we may have succeeded more than we have dared to think. Anyway, here are some comments for what they are worth. Obviously, the list is incomplete, but they are the observations and recommendations which came most readily to mind this week.

A. FACULTY

We have been slow in working out our own identity. We have been hesitant to work out this identity because we have not wanted to separate ourselves from the students. I have consistently felt it imperative that we build a community of trust and openness between us that will enable us to see ourselves clearly through honest feedback and to develop trust in our colleagues through confrontation. The sense of identity developed will not separate us from the students but will enable us to form the basis of genuine encounter with them. Such identity does not mean an escape into a role. It forms the basis of self-determination.

Observations:

1. We have lacked a support system for faculty, although the T-Group has fulfilled some of this need. How does a member of the faculty find support through some trial, either by way of a threat to his employment or to his sense of frustration in his teaching? Or, how does he experience support for the pursuit of a goal which may involve a risk?

2. We have tended to overreact to various situations. They have usually involved responses made more as the result of the intensity of commitment rather than any specifically hostile response. We have tended to be protective of the college as a new organism which we are nurturing.

3. We have had the problem of an overload both in terms of the intellectual and the emotional demands made upon us. The effect of the overload has been evident in defective decision making. The organizational demands may be less this next year, but the emotional demands will continue to be heavy. The greater the flexibility in a system, the greater the emotional load. The straight line method of management authority and decision making provides security as well as rigidity. When the clear-cut lines of authority are not present, decision making is both more chaotic and emotionally demanding.

4. We have not developed experimental methods of tracking and learning. While we have not had time to be reflective about what we are about, our lack of experimenting is also the result of inadequate skills.

Recommendations:

1. Training of faculty as facilitators of both interpersonal encounter and of the learning process through small groups is a priority of the interpersonal program. It will begin with a period of training in August under the direction of Kouji Nakata, and will continue through the year as he makes his consultation available.

2. Sections or subgroups within the faculty need to meet together regularly to develop an intimate support system. A small group meeting around a particular interest will enable them to develop skills of:

   (a) commitment to a specific educational goal within the more comprehensive goal;

(b) accountability to one another in both personal and educational goals; and

(c) challenging colleagues and students to persist in setting and achieving goals of personal and academic growth. I have already proposed that those of us with special interests in interpersonal and affective education meet weekly to build an intimate relationship and to develop the skills necessary to achieve our goals at Johnston College.

These groups could request a consultant to act as trainer or process observer.

3. There is a need for a consultant to provide ways in which the separate sections of the faculty can be brought together–to provide interfacing between the groups, and to provide a way in which the faculty as a whole can function most effectively in decision making.

B. COMMUNITY MEETINGS

Community meetings have often become the source of frustration, indecision and self-defeat. Faculty and students alike have tended to evade responsibility for this important part of community life by passive withdrawal. We have tended to perceive a meeting as a giant "tantrum mat" upon which we stamp our feet, rather than seeing it as a potential learning situation.

Observations:

1. We need to refer items involving decision making to task groups which could either be given the power to implement an action or to report back to the community.

2. The faculty needs to volunteer as process observers–or to train process observers in at least political and interpersonal areas of community decision making. The faculty should become much more involved in the *process* of the meetings, enabling them to become learning situations.

3. We need to study the emergence of power blocs and learn how to cope with them.

4. We need to help provide some time limits of the meetings so they do not continually interfere with regularly scheduled seminars.

Recommendations:

I propose a seminar available for both faculty and students for training both as chairmen and as process observers of community meetings. I think Kouji and I could serve this function. This seminar could study the *process* of the community meetings and present their observations either as process observers or as chairmen.

C. ACADEMIC ACHIEVEMENT

In comparison with the usual college, it seems to me that our failures have been more gross, and our successes more dramatic. This result may be expected in an experimenting college. In general, I think our production was low, especially in terms of written work. I suspect that the amount of reading was above average for a freshman year of college. However, some students read very little and others read considerably more than would normally be expected.

Observations:

1. Some students have not been able to function well with the lack of structure. They should be confronted with their inadequate functioning and assisted in either learning how to use this system or transfer to a college where they can function.
2. There seems to be a general lack of a sense of learning together or of teaching each other in a community of learning.
3. There is a general lack of understanding about contracts and evaluations.
4. There seems to be a need for some type of mid-semester evaluation of seminars and tutorials.
5. The percentage of students able to use this free system constructively needs to be increased. Perhaps special assistance could be provided for some students who recognize the need to develop their learning skills.
6. I think that returning students can become a very important resource to new students in assisting in the learning process, setting up community involvement, and in the teaching of subject matter. Their teaching assistance needs careful planning and supervision.
7. I don't think we have challenged some students sufficiently to high expectations of themselves. Some are able to set such goals by themselves, but others are not realistic in appraising their potential.
8. Some guidelines in negotiating a contract could be helpful to students.

Recommendations:

1. I propose that increased attention be given to the building of a contract, and I intend to include an introduction to contract negotiation and evaluation during the orientation sessions.
2. Some introduction to the process of learning will be included in the orientation session, and perhaps a series of workshops could be provided during the semester.
3. I propose that the faculty set out some guidelines and dates for contract negotiations, mid-semester evaluation and final evaluations.
4. Twenty students are being invited to participate with faculty in the training period August 27-September 2. Hopefully, these students will be able to use their training as facilitators of learning with new students throughout the year.

D. ORGANIZATIONAL STRUCTURE

Since there is some structural ambiguity in the relationship of Johnston College to the University, it is especially important that we work out our organizational problems carefully. We have a new president, and we need to give him time to work out a more satisfactory structure. It is imperative that our organizational relations facilitate rather than impede the realization of our goals.

Observations:

1. **A better relationship between the University administration and Board of Trustees and Johnston College needs to be worked out immediately.**
2. **The Board of Overseers needs to have a closer working relationship with both faculty and students.**
3. **Lines of authority are not clear.**
4. **Responsibility for areas of faculty power and responsibility are not clear.**
5. **How can we have increased organizational clarity and efficiency without the corresponding rigidity?**

Recommendations:

1. **I propose that with the consent of Dwayne Orton, a consultant be secured to serve as a process observer in the planned meeting of the Board of Overseers with students and with faculty.**
2. **I recommend that Kouji Nakata serve as process observer at faculty meetings and help facilitate the development of an effective structure without the attending rigidity.**

E. MINORITY STUDENTS

We have not solved the problem of institutional racism which is an unconscious part of most institutions dominated by the white majority.

Observations:

1. **Many black students have not related to the community as a whole. They do not feel heard, and they are not ready to relate openly.**
2. **The black encounter group initiated the last semester seems to have fulfilled a need for black students to meet together, and they want to continue this group next year.**
3. **The brown students seem to have worked out their relationship with themselves and with certain members of the faculty and staff, especially with Mike Tirado and Lino Lopez. They have had assistance in relating to each other and to the community in action concerns.**

Recommendations:

1. A black encounter group is planned for this next year to include the returning students and others as desired.

2. I will propose that new black students constitute 50 percent or more of a mixed group in the opening orientation and training sessions.

F. THE COLLEGE COMMUNITY

I have mixed feelings about the state of the community as a whole. Although there are some serious problems which we encountered, most of them are normal for any college, and particularly normal for any first year of a college. These problems need to be taken seriously, but I am optimistic about the health of the community.

Observations:

1. Students, as well as faculty, have experienced an emotional overload. The lack of structure and definite basis of decision making has loaded students with a great deal of anxiety. They do not have some of the normal bases of security such as a sharp line of demarcation between faculty and student responsibility and grades.

2. Students, as well as faculty, have not had an adequate support system to provide for maximum realization of potential. They have not had the support necessary to give them the security to take the risks necessary to their growth.

3. Students have reacted against the T-Group design for several reasons: (a) It has become the focus for a requirement for graduation and they are reacting against any requirement. (b) The T-Group design is too limited for the purposes of an educational institution. They need a variety of designs to facilitate personal growth. (c) Although there were at least 18 different personal growth groups functioning during the year, many students did not have any continuing experience of personal growth. Some of these students felt that the trust building initiated at Pilgrim Pines had been a betrayal. (d) Some students are able to experience personal growth through programs of service to the community, and personal growth can occur in situations other than small group experiences.

4. The community takes itself too seriously. One mark of maturity is a sense of objectivity, or a sense of humor. We need to be able to look at ourselves and laugh at ourselves in our most serious intentions.

5. The total design needs to provide some structure whereby students can experience psychological distance as well as intimacy.

6. Accuracy and speed of communication is more important than in a usual college because lack of clear-cut structure increases possibilities of misunderstanding.

7. I have seen significant positive changes in many students and feel that for the most part, the mental health of the community is superior to that of the usual college.

8. We need the maturity which upper-class students bring to a campus, and hopefully some of these needs will be met next year.

Recommendations:

1. I will propose a variety of programs of personal growth to be offered to students, and these programs will be continued throughout the year. We will offer a sufficient number of programs to enable each student to have at least one semester of a personal growth experience each year.

2. I propose that returning students and faculty take an active part in planning at least some social events such as the Medieval Festival.

3. I propose that the Associate Dean of Students develop something like "dorm families" in order to provide a support system for students in their living areas.

Director, Interpersonal Dimension

Dear Pres:

The best I can do at this time for an annual report. We tried, never feared to try, all sorts of new ways of teaching and learning as fully humanly as possible. For psychological neophytes like me, this was amateurish because of general ignorance of a great deal of group and interpersonal processes. Common sense helped a lot, but I needed more sensitization to my own psychological hang-ups that in turn caused a smooth class to stumble when deprived of what we agree is artificial structure (lectures, tests, formality, etc.).

The GREATEST things are heart soaring: (1) we have the encounter-group attitude permeating this whole place to get us to bring joys and disappointments out openly, a safety valve, a keeping together, and (2) we have a continuing mandate to keep trying better, new things even when we have not succeeded completely with some recent experiment. Overall we have this great vision of striving toward fully human education, in which I can work for truly human science. That's the greatest. That's where I came in last year and it's where I am now.

This has been undoubtedly the most profound year I have ever spent. To think it is still two weeks short of being a year since I first heard the words "Johnston College." Incredible. Spiritual.

Fellow in Chemistry

Dear Pres:

THE OUTSTANDING ELEMENT in my year here, so far as I am concerned, relates to the crucial impact of sensitivity training upon me as a person as well as upon my classes. Above all, it made possible the kind of class environment I have always sought in my teaching, i.e., one that is unhindered by the barriers that inevitably attend traditional academic pomp and circumstance. Accordingly, I feel that I have been able to reach my students as never before at levels that are crucial for development. Since I came to

Johnston precisely under the intrigue of the introduction of the affective levels into pedagogy, I consider myself richly rewarded by the year's experiences. To be sure, I have never worked at anything so intensely as I have teaching at Johnston this year. But, never have I been rewarded quite so inordinately. I could do it year-round!

It may sound odd, but I have kept abreast of my field (indeed, even expanded interests and competence) as never before. I know the reason for this: release from the endless round of lecturing delivers one over into the joyous world that is scholarship. Teaching through seminar and tutorial demands more scholarship than lecturing, for one must be prepared to engage all elements of one's field at any moment (without notes and manuscript scrupulously prepared beforehand!) given the demands of the seminar and student need. That has been a real lesson, and one that I have been pleased to learn.

I wouldn't trade this year at Johnston for all the grants at UCLA or Cal! I consider myself a very fortunate guy.

Fellow in Religion

Dear Pres:

Herewith a year-end report for the momentous year 1969-70. For me it has been without question a substantial and rewarding experience. I have appreciated, and taken advantage of, the opportunity to experiment in academic style and content, with the desire of moving closer to the academic needs of all those concerned. Some of these experiments have been unsuccessful, mainly, I think, when there was too loose a definition of points of departure, insufficient mobilization of group involvement and inadequate support materials. Although lack of faculty expertise has sometimes been a further factor, I do not believe that this is necessarily dysfunctional; my Gandhi course proved remarkably rewarding for most of us, even though we all had very little prior knowledge to go on.

Rather than go into specific experiences or recommendations I will limit this report to outlining basic individual and group needs which I have experienced as the year proceeded.

In academic content there has been a pronounced desire to center on exploring human nature and personality—as evinced in the specific examples around us—in order to become closer to people as people. Under this approach people themselves become the primary object of study, rather than ideas, theories or events as presented through written materials. Needless to say, our program gives plenty of emphasis to this concern through the Interpersonal Dimension; but we may as well recognize that this need cannot—and in fact should not—be limited to courses of a recognizably psychological description. It applies wherever man is the subject of study, and man rightly *is* the subject of study virtually throughout our program.

One might go further and say that the desire to explore human nature is sufficiently strong, that it should not be limited in techniques and strategies, any more than in dimension or disciplines. The resistance which we encountered towards some (not all) of our sensitivity training experiences seems to me attributable to feelings of their inadequacy, and not to opposition to their underlying objectives. The ends hold good, and it is up to us to evolve viable means which defer to those ends.

In interpersonal attitudes there appears an ingrained need to struggle with the blocks implanted by authoritarian education. These blocks seem far more widespread than I, for one, anticipated. They are involved wherever external authority, rather than the immediate encounter, directs the conduct of human relations. This is to criticize authority, not as an intrinsic force, but as and when it impedes open exploration of our human environment, which it does all too frequently. I would like to see us affirming, or at least encouraging, assertive, demonstrative behavior as much as possible, even at some risk to our personal equanimity. Met with open response this behavior and interaction seems more conducive in the long run to social respect and affection than behavior ruled by authoritarianism. As a principle this is in fact obvious; it is in practice and implementation that it tends to run aground.

As an example of this propensity, I myself tended this last year to lean over backwards not to be too assertive in class because of assuming that assertiveness in a "teacher" can too easily have the impact of authoritarianism. In some cases (e.g., my participation in the Current Events seminar) I found that it had precisely that impact.

The point is that we have each and all to learn to distinguish between assertiveness based on role or status and assertiveness based on individual definition. It seems to me that the latter is really an attempt to become more open and, paradoxically, again to move closer to others. Its "Here I stand; this is me" character, though often negative in expression, is surely an authentic component in the process of finding and establishing oneself through interaction with others. In short, I see this process as another means of expressing the same basic desire to explore and move closer to people as people and believe that it should be tolerated and affirmed.

In relation to general conditions there is a powerful need to find reality in the midst of the doubts, anxieties, appearances, hypocrisies and straight confusions which beset our time. Isn't this what all this talk of relevance boils down to? The catch comes in defining reality. We have had a good deal of talk in the last year about the need to face "reality," but there have been widely divergent (although generally inarticulated) assumptions as to what, or whose, reality has to be faced.

This may be the nub of the problem which we will have to deal with. If we associate reality with material and institutional power, then we will surely have a power struggle on our hands for the control of that "reality." Rather than take this line I would urge that we use our academic independence to seek out and substantiate less exclusive notions of reality, beginning with study of our own individual and group participation in the total nexus of whatever reality is. This would demonstrate our intention to move away from authoritarianism (which underlies exclusive notions of reality) and towards the notion of life as an interlocking rhythm. We are already doing this, but we have been repeatedly deflected, and we need to be sure to retain it as our sustaining objective.

Again I am sure that this pluralistic approach is in the long run less conducive to antagonism and more conducive to social and personal respect, including for those who may not share the same objectives. Having myself sometimes followed a rather militant line last year (at least vis-a-vis some of those in authority), I intend next year to try pursuing a more inclusive (if not less assertive) strategy for fulfilling this need.

All these needs are subsumed in your educational premises for Johnston College. Our experience last year seems to me to reinforce their validity and to suggest that we should certainly continue to maintain them as our primary objectives. Our anxieties and mistakes prove that it isn't at all easy to implement these objectives and emphasize that we need more time and experience with which to develop a feel for their implications in actual conduct. This will necessitate further tolerance from those in authority, and I trust that it will continue to be made available.

Director, International Dimension

Dear Pres:

I wish to enumerate some of the more important items of learning I accomplished during the past year.

A. In my area of academic pursuit, that is, sociology and social anthropology as they particularly pertain to the understanding of organizational behavior, freshman students can be confronted with materials for their reasonably effective comprehension, which would normally be expected of sociology-anthropology major upper division students or of graduate students. This possibility was the most unexpected discovery on my part, and it is contingent, undoubtedly, on the kind of learning processes we were able to provide at Johnston College.

B. In spite of the above positive promise demonstrated by our style of teaching-learning, it was also true with most of the students I encountered in the successive classes and elsewhere, that the amount of transfer of learning and effective application of concepts and methods of analysis they thought they had learned, to varied ongoing life situations, is rather limited. Having made this statement, however, I do not wish it to have any comparative efficacy with other, more traditional sorts of learning situations. We just do not have any objective, meaningful way of comparison. What I can say is that in a traditional setting, I would not perhaps care if there was any amount of transfer of learning.

C. I have taken the meaning of "education for relevance" more seriously than its publicity value for Johnston College. I have taken it seriously because of my initial reservation about its implication for education of the ephemeral. I have observed directly and indirectly this negative aspect of the concept, particularly in issue-oriented seminars at Johnston College and elsewhere, where there is a danger of engagement with polemics giving students a pretension of learning. I have thought that we should not be concerned with the training of sophists with the passing knowledge of *doxa* of the century. Other universities as well as political groups of all kinds from SDS to the American Legion are available to do just that.

Through the experience of this past year, I have come to realize more keenly that what we are trying to do here is to stress the *relational* implication of the notion of relevance when students try to identify their needs and objectives for learning. What I mean is that we help students define their learning in terms of their sense of communal responsibility–relatedness of himself with others, and thus eventually his sense of worth of himself in terms of his responsiveness to others.

I have come to this realization through my discovery that those students who expressed, and exhibited in my eyes, satisfaction in their learning tended to be those who made a real effort to learn in a communal context–in an active reciprocal context with other students as well as with professors. Conversely, those students who either shunned away from articulating their sense of communal responsibility or refused to accept it as a significant criterion of relevancy, tended to be limited in their learning in spite of a higher abstract sophistication on the surface in terms of the depth of understanding of themselves and the breadth of their perceptual range. Doing their own thing simply did not give them the reward they could have expected; and most damagingly their ability to formulate problems remained relatively uncultivated. They remain more vulnerable to the demonic power of intellectual chauvinism as well as to that of antiintellectual nihilism.

The meaning of education for relevancy at Johnston College may be interpreted as our attempt to restore humanism in the educational process. But I believe that if by this humanism we mean an introduction of the learning-by-doing Americanism into higher education, our fate is destined. The world is too tragic and too cynical to accept that type of humanism any more. We need a more realistic direction-giving, commitment-giving ethos than the rainbow "humanism" of the pre-1940 Western antiquity. At this tentative moment, I can only see a glimpse of a direction we seek to encourage in the Johnston College program, which I express as our sense of communal responsibility, or an existential awareness of man in communion with another man as an essential basis of education. Trust we often speak of is indeed an aspect of such human reciprocity. The openness we have cultivated is a prerequisite of reciprocity without demanding necessarily compulsive compliance with a uniform set of values or casuistic morals.

D. It has become abundantly evident to me that a major preoccupation of our freshman students is in the area of the interpersonal, not intercultural, international nor environmental. This is no surprise. I suspect that the proposal for the QFM seminar last year took into account this obvious socio-psychological state of our young people. I reiterate the obvious, for I think that it should somehow be reflected in our curriculum strategy and in the advisement of individual students.

To "prefabricate" and structure their preoccupation in such a form as QFM seminars is meaningful to only a segment of the students. This has been the basis of an ironical frustration we experienced in the Quest for Meaning attempt. What is more necessary is our own increased awareness of the nature of their preoccupation in the context of our helping them to formulate needs and problems for individual seminars and tutorials.

E. I entered into the sensitivity training with a mixed feeling of curiosity and reservation; reservation in the sense partly of its implicit assumptions and cultured values that are not altogether congruent with my value system. I believe I have learned to overcome much of the hurdle I felt initially; and I am more positively convinced of its effectiveness in the kind of environment in which Johnston College education takes place. I feel strongly that the affective education is strategically important for the enhancement of the basis of trust and communal responsibility. I believe that in spite of the importance of initial and periodic programs we provide for ourselves and students through T-groups and so forth, nothing replaces, or is more important than, its integrated place in day-to-day routine interactions in seminars and tutorials.

F. There are a few matters relative to student evaluation that continue to hover over me. The first concerns the problems related to retention. If we are to be consistent in our experiment in terms of its basic assumptions about education and the consequent structure and mechanism of day-to-day operation, it seems obvious to me that the advisability of student withdrawal from the College must first be recognized and initiated by students themselves. There are always "misfits" among those who should not continue at Johnston College, who do not even come to recognize this. The number of such "misfits," as we already know from our experience, is very small. A mechanism required to take care of this small number of students, who should or could have been identified initially in the admissions program, should never become the procedure to be followed by all the other students. My assumption here of insisting on consistency of our basic premises of education does not lessen our responsibility to tell students to leave at any point in their career, if and when they prove to be unmistakable "misfits" at Johnston College. I am concerned that we should not devise a procedure in the name of education at Johnston College where we unwittingly take away from the majority of our students their primary responsibility for initiative in their own education.

The second point has to do with an element of anxiety expressed by some of us as well as by some Overseers concerning a "standard" to "measure" the degree of personal growth of students, which, without any question, is an essential phase of our education at Johnston College. I am appalled by an implication of this anxiety or concern I sense: that there exists in this world such a thing as an objective, scientific, and therefore educationally meaningful standard to measure the degree of individual personal growth which can be used, presumably, for determination of student retention. I only hope that I am mistaken in my inference, or that I am overreacting to the expressions of a few. Indeed, serious psychologists would be the first ones to disclaim such a notion. My sense of dismay is compounded by a realization that it is we academicians, particularly those in Social Sciences, who have perhaps unwittingly but successfully brainwashed laymen and educators through the aid of educational technocrats to the point of their embracing such an illusion. One does not need to know all about the hidden Lockean assumptions and the games of positivism behind the advancement(?) of American psychology to realize the dangerous magnitude of this notion. One does not need to know, perhaps, that there is no logical or scientific way in which he can identify the statistical measures of validity and reliability of assessing personal growth with their educational legitimacy for individual persons. One does need to be reminded, however, that personal growth is by definition subjective par excellence. It is so in spite of the fact that psychology helps us to define and partially explain certain dimensions of personality and growth patterns.

I am in no way advocating that at Johnston College we should not evaluate the extent of personal growth of individual students, or that we have no way to evaluate. We already have a mechanism for evaluating personal growth that is built into the contractual procedure. What has not been worked out systematically, perhaps, is the identification of the range of assumptions and values used by more than two hundred evaluators in the College for their subjective assessment of

> subjective personal growth. Some of us, who have been directly involved in the development of the intercultural program, have begun to do this task for our purposes. In spite of our frequent use of the most unimaginative jargons of the hydraulic input-output model of communication, and in spite of our preoccupation with the cybernetic constipation with too many problems-to-be-solved, I only hope that at Johnston College we refuse to become the prey of technological imperialism under the illusion of being scientific. That has been one of the most impulsive Janus-faced demons we have met since the days of Adam and Eve.

As we look forward to the new academic year, there are two things that seem to be of critical importance for the future of the College. The problem of financial support is too obvious to mention, and my idea about it is no better than others.

The first is the problem of size. From my experience elsewhere, but particularly at International Christian University where it was originally projected that the undergraduate division be kept at 600, the size of 300 next year and ultimately 600 by 1972 is a reasonable, modest one. In terms of the financial pressure we have, this rate of increase is perhaps unbearably slow. However, while we are at an experimental stage of this experimenting school, the rate is unbearably fast. The point is that the Johnston experiment does not depend on the student-faculty ratio of ten to one. We are discovering that it depends critically on the relative proportion of students an individual faculty member gets to know as individuals. It depends on a simple reality that openness, trust, and other related operational values which are the keys to the College's success as well as to the satisfaction of individual participants (including the donors) cannot be institutionalized. They can be enhanced or discouraged through certain structural arrangements, some of which we call "the interpersonal programs." But we have no way to institute them by their very nature. They are humanly created and personally transmitted. Those who have been nurtured in the Judeo-Christian tradition should be more keenly aware of this human condition of human values than others. Whatever the circumstances may be that have necessitated the coming enrollment at Johnston College, we as well as the controlling Boards must be clearly aware of the fact that we have built into our program next year a factor which will contribute more *negatively* than we care to admit toward the success of the experiment.

The second is the assumptions and policies related to our governance. We can reasonably anticipate next year increased pressures for confrontations and politicization with an increase both in quality and quantity in the internal constituents of the College. Because of the nature of these pressures, which stems from the changed conditions of society and the increasingly professionalized scholarly world, the prevailing mode of governance as currently conceived, on the one hand, by the controlling Boards, is sadly inappropriate and ineffective in processing these pressures.

On the other hand, we are internally in the process of groping to develop workable arrangements for distributing and sharing power that are consistent with the objectives and the emerging ethos of the College. In this context, Chancellor McCoy's introduction into the assumptions of our government the concept of expertise is a viable one. However, even this concept has definite structural and functional limitations. In fact at some important points the concept is at variance with the Chancellor's authority as it exclusively derives from his relationship to the controlling Boards. We will certainly continue in our effort to work out some satisfactory modus vivendi to handle the educational problems of the latter part of the 20th century under the office of the Chancellor, who must, in turn, respond to the requirements of the statutory rigidity and ideological framework relevant to the pre-19th century West. This is a vain effort, however, so long as controlling boards refuse to understand the implications of this reality gap, and to recognize that by some means the authority relationships necessary for college governance must be *legitimized* in the collegiate setting for them to be authoritative.

It is the immediacy of increased disruptive pressures I sense and the hope I have of making education relevant to young people at Johnston that I am compelled to register these concerns.

Fellow, Intercultural Dimension

Conclusion

The scope and rapid pace of change in our postindustrial age demands that we help individuals develop flexibility and openness in their life styles while sharpening critical faculties in problem solving and encouraging creative power in dealing with ideas and persons. Increasingly, individuals will be living and working in a variety of cultures due to the growing interdependence of political and economic systems. In such a milieu, any person, to be effective must possess a certain certitude about his own identity and those values which form the basis for personal attitudes and decision making. Yet, he must be appreciatively open to differences in other persons, cultures, and value systems as a prerequisite to working well with them.

A healthy spirit of self-criticism, an awareness of process in personal and group interaction, combined with imaginative use of problem-solving methods will strengthen man's chances for survival. Our technological age is in great need of philosophical maturity and keener aesthetic awareness. Developments in the biological sciences generally, and in genetics particularly, will crucially test man's wisdom. All methods of the academic disciplines must be applied in interdisciplinary fashion if such problems as ecological destruction and international conflict are to be solved or contained. In its program design and methods, Johnston College is attempting to help students prepare to meet such future challenges. Evaluative processes will play a key role in helping us determine the relative success of our efforts. Results of this continuing experiment will be shared with the educational community.

Chapter 4

FAIRHAVEN COLLEGE

Charles Harwood[1]

Traditionally, formal evaluations of colleges have limited themselves to a single approach. Thus, a researcher may employ tests of knowledge and attitude and observe changes in students during the college years. However, the vital audiences which furnish potential students, potential faculty, and who give support to the college may well ignore that data and judge quality in terms of things as vague as "general reputation," or reputation as based on information presented by public relations-minded college spokesmen. Thus, formal evaluations often result in published compilations of facts and figures, whereas the evaluation that makes a difference to a college is done by the audiences of consequence with no heed to the data. The problem is confounded by the different "languages" used by professional evaluators and these audiences. There exists the general question of meaningfulness. Part of the answer has been suggested by Henry Murray in his pioneer work on the study of personality of college-age men. Murray (1938) noted that "all proposed formulations must be supplemented by the language of common speech [p. 206]." Changes in the goals of education and the different interests of various audiences add to the problem faced by the evaluator.

Therefore, rather than propose only one approach to evaluation of Fairhaven College, three strategies are to be involved in order that the evaluation may prove useful to the various audiences involved over a long period of time. The first, and unique approach suggested here, is that of paying attention to the audiences of consequence and the sort of information upon which they base their judgments. In the past, this aspect of evaluation has been accomplished largely by the brochures, rumors, news stories, press releases, etc., which have been circulated. Although this first strategy is not usually included in formal plans for evaluation, the fact that it perhaps is the most vital aspect of evaluation means that it will be considered. However, because of the limitations of evaluating a college by "reputation" it is not proposed as the major strategy but rather as a guide for the development of other approaches.

---

[1] Editor's note: Charles Harwood died before completing this essay. Except for very minor editing, his draft is here reproduced. Paul Woodring graciously agreed to add some concluding remarks about tentative evaluation plans at Fairhaven.

The second approach is also relatively uncommon. This is the long-range study of what happens to the individual during his life and the relation of his life experiences to his college experiences. One example is provided by Terman's study of genius in which a large group of bright children were identified and studied for many years and the correlations of high intelligence and many aspects of their lives have been reported. The Terman example is not particularly concerned with the impact of college. Newcomb's study (1967) of Bennington College students after 25 years is a specific example of this approach. Longitudinal studies of the lives of normal individuals also provide examples of the relating of college experiences to later life. Specific examples here would be the type of case history reported by R. W. White in his *Lives in Progress* (1955). Sanford, in his major work, *The American College* (1962), states: "In the study of organizations, as in the study of individual personalities, the intensive investigation of the single case remains the best means for revealing the wholeness of the system in action [p. 7]." Sanford here is referring to the chapter in that book by Jencks and Riesman (1968) and their use of case studies of particular colleges as a means of illustrating major points. Similarly, it is here proposed that case studies of individuals over relatively long periods of time which relate college experiences to their behavior in later life will be a particularly valuable approach. In the evaluation of Fairhaven, the long-range approach is to be the major aspect of the evaluation.

The third strategy, which for purposes of naming will be described as "face validity" approach, includes a variety of commonly employed means of evaluation. In addition, at least one which rarely has been systematically employed before will be included. The latter is the systematic observation of student behavior which appears to be appropriate and important to the educational process. An example of this would be observation of student discussion topics outside of class to ascertain whether or not the behavior was consonant with the educational goals of the college. Continued discussion of material related to educational goals would be scored as serving an educational goal. The more traditional of these approaches are exemplified by use of the Omnibus Personality Inventory (Heist et al., 1969), the Institutional Functioning Inventory (Educational Testing Service, 1968), the CUES (Educational Testing Service, 1969), and tests of academic achievement. Each of these would be used and examined in terms of the goals of the college. The development of the student during his school years will be included under this third strategy and would form the natural beginning of the long-term study mentioned above. The "case history" method of study also describes this approach. Again, to employ some of Murray's (1938) insights, "case histories are the proof of the pudding." He also emphasized the wisdom in having a group of experimenters use the same group of subjects in order that different scholars and different theories can be exposed to the benefits of comparison.

The rationale for this multistrategy approach needs further elaboration. One has only to examine Feldman and Newcomb's two-volume report of *The Impact of College on Students* (1969) to grasp the magnitude of studies to date, but has only to discuss colleges with "informed" colleagues to discover how seldom they are considered in rating a school. This recent and inclusive review provides a most important source of information about the impact of college and will be frequently referred to in this paper. However, setting aside a small group of individuals whose careers are involved in research in higher education, higher educators usually employ other information as a basis for their conclusions. Similarly, potential faculty, students, and their parents do not usually base their judgments on the kind of data included in this book. And to a large degree, neither do the high school counselors, legislators, editors, and people in general. It is here contended that unless evaluations take into account the kind of audiences for whom they are intended, the information will prove irrelevant. Paul Woodring has noted that evaluation can be described as "subjective judgments of people (who) count." In addition, as a perusal of the Feldman and Newcomb volumes will attest, there is a good deal of variation in educational goals over time. Much that was considered of import at a particular moment now is judged trivial and, no doubt, much that fills our headlines today will also soon be insignificant. Because we are unable to know the future and yet need to be collecting evidence now which may help future judgments, it is vital that an evaluative technique that allows the casting of a wide net be provided. At a recent meeting of leaders in innovative higher education, the author was much impressed by the *lack* of interest in research reports concerning the personalities of college students. Because the data being presented were some of the best available and the audience was composed of leaders in experimental and innovative colleges, the urgency of new approaches to the matter of evaluation was further emphasized. If specialists in the development of residential, liberal arts colleges are not interested in the usual approaches to attitude change, it is highly unlikely that other audiences will have such interest.

It is *not* proposed that the first strategy, that is, the interest of the ultimate audiences, be directly attempted. This approach, if pursued directly, could change the model from that of the scientific researcher to that of the advertising man. However, this strategy does suggest vital guidelines. Therefore, the best in scientific measurement must occur to the extent that information which has relevance to some present or future audience can be obtained and presented in the language of these audiences. The difficulty of defining "liberal education" plus the changes of educational goals over time further emphasizes the concern for a catholic and comprehensive approach to evaluation. Finally, it should be remembered that, in the ultimate analysis, evaluation is a matter of human judgment. Strategy number one reminds us that those who make the final judgment cannot be ignored simply out of deference to scientific convenience and methodological rigor.

Liberal education is concerned with the development of an individual in relation to his future life. Therefore, it is mandatory that a proper evaluation take into account the relationship between educational experiences and this latter life. Just as most would agree that mastery of a particular body of data for a particular examination is not necessarily of consequence to one's later life, it follows that any study limited entirely to specific behaviors during college years cannot demonstrate correlations which may obtain between college experiences and later life. Granted that the mere demonstration of a correlation between certain experiences and later life does not prove a causal relation, nevertheless it is a necessary first step. Terman's study of genius has proved to be especially valuable, although it is a longitudinal or correlational study rather than an experimental one. Study of the relation of particular educational experiences in the undergraduate college years of life with events of later life should also prove of value. Hence, the identification of college experiences both in and out of the classroom and the investigation of their relation to later life is the key task of the evaluator as well as the planner of higher education. Enough instances of the sort reported by R. W. White in his case history of "Hartly Hale" (1955) would be strong evidence for the importance of out-of-class contacts during college. In that history, one of Hale's argumentative dorm mates played a crucial role in his decision to start his medical career. As Hale reported this important incident:

> One night for some reason or other, this was out of the clear blue sky, I don't know why or what brought it up, he decided, he got it in his mind, that I should be a doctor instead of going into advertising. And so he started to argue with me about it, and when morning came I agreed that he was right. He did a very logical, very thorough job on me, and so I went down the next day and I changed my field of concentration to biology. Then I called up my family and told them what I'd decided, which they didn't like [p. 47].

Thus, we have a convincing example of the possible importance of a residential requirement as a means of providing important aspects of an educational environment. Hale's medical career, by the way, was a distinguished one. The case history also provides examples of how certain personalities of professors interacted with that of the student and the opportunity to become involved in a research project (in the honors program for which he was not qualified according to his grades) which furthered his development as a student of medicine.

Sanford (1962) vividly describes observation of alumni in their home situation which tells of the boredom in the lives of educated Americans plus the lack of civic involvement. In visits to ex-students in home situations,

> It was explained that in this town there was nothing for educated people to do. There was a symphony concert one evening a week, but for the rest of the time people engaged in a round of parties, with heavy drinking the rule. With many of the women, the drinking started with three cocktails before lunch with the girls downtown. Then, about 3:00 one of the group would say, "I have a wonderful idea—let's all go over to my house and have a drink"; this they did, and thus whiled away the time. . . . Undoubtedly the women of this social group, all of whom were college educated, had had many occasions to complain about the lack of stimulation in this culturally barren and sociologically depressed community. Depressed is the word. The center of the city revealed no signs of civic pride; there were shabby public buildings and a continuous snarl of traffic. In going any distance from the center of the city, one passed through vast slums. The delinquency rate matched that of any of our large cities—and the college educated people of the city were bored and could find nothing to do with themselves. Where is the capacity to go on learning and to go on demanding the kinds of aesthetic and intellectual satisfactions that were enjoyed in college? Where is the sense of social responsibility that our colleges seek to encourage? In the case of the women, where is the sense of independence and self-confidence that would make it possible for them to assume leadership in a desperate situation? For that matter, where is the sensitivity and social awareness that would tell these women and men that the situation is desperate?

A wide variety of college experiences both in and out of the classroom is judged by various students of higher education to be valuable as part of a liberal education. These judgments are made on the basis of the goals of the specific evaluator and his guess as to how certain college activities will affect the student's later life. The activities themselves range from mastery of specified subject matter, development of particular study, writing, or research skills, living experiences in which cooperation, group government, living according to a particular code of rules, and exercises in certain kinds of self-discipline are involved. Although what may be on the face of it an important educational experience from the perspective of one educator may be judged as being of little consequence by another, there are many aspects of college which have frequently been identified as important for education. A wide sampling of these is required by the third strategy which can have a shorthand title of "face validity." Some of them are defined in terms of changes in the behavior of the student whereas others are defined in terms of experiences which the student meets. Thus, for the traditional mastery of subject-matter goal, the student's performance on tests can be taken into account. Others are content to measure the success of a program in liberal education by examination of the curriculum, identification of the number of "great books" or disciplines studied, and the judgment of the value of a curriculum on the basis of this examination. Recently, great emphasis has been given to the personality development of the student. One aspect is exemplified in the long-term studies of Newcomb (1967) and his students who have examined the change in political thinking during the college years and then, pioneering in the long-term approach described above, studied these students again 25 years after they have left college. Newcomb, like many other scholars, uses such devices as the Omnibus Personality Inventory and the CUES to measure the kinds of student personality and the kinds of environment which the student faces. These various tests allow evaluators to, in some instances, look for change in student behavior and in others to learn of conditions which the evaluator believes to be important for the student's education.

The behavioral observation approach follows the same general logic in that certain student behaviors are judged conducive to the development of a liberally educated person and, therefore, if they can be observed, a favorable evaluation can be made. Thus, a frequently cited goal of living-learning colleges of "continuation of discussion outside of the classroom" can be measured by literally listening to what the student discusses outside of class. Similarly, educational goals in the realm of social action can be measured by scoring the percent of students who are taking part in social or civil action, college governance, residence hall governance, etc. Concern for one's fellow man might be shown by everyday indication of trust–doors left unlocked, books left in public study areas–and by the consequences of this trust; that is, relatively high or low loss rate.

A Composite Evaluation Plan

A wide variety of evaluative activities during this student's college years makes up the first major step. In their written evaluation of their tutees (all students have a faculty tutor), each faculty tutor will be encouraged to identify experiences which are or have the potential for being significant in the student's life. Faculty are encouraged to ask the student to make similar analyses and keep a log of events. A letter of evaluation by each faculty tutor is mandatory whenever a student leaves college to serve in lieu of grades. For research purposes, this letter will be prepared under guidance of the evaluators and will be done annually. The students in the research group will be asked to keep a log of events which they believe may prove to have important bearing on their later life. Guides to log keeping will also be furnished to each student to ensure coverage of many areas.

More traditional evaluative strategies will further broaden the data collected and will also serve to sensitize faculty and students to areas to which they should pay special attention. Thus, for example, the fact that Fairhaven students score high on "autonomy" may lead a particular student and faculty tutor to be sensitive to the way in which participation in college governance or social action during the college years may serve in the development of a student. On the CUES and on the Institutional Functioning Inventory, Fairhaven is already known to be particularly supportive of the individual student as a franchised member of the college community. This political status may lead to college experiences which set trends for his later political-social life.

Routine evaluation which is already underway at Fairhaven includes the following: the Omnibus Personality Inventory, the Institutional Functioning Inventory (administered to the faculty), the student's high school and college grade point averages, the background of students as measured by an American Council on Education questionnaire, and a locally developed questionnaire which measures student attitudes towards Fairhaven. Previous exploratory work has employed the College Characteristics Index (Stern), and use of the CUES for further examination of the particular kind of environment at Fairhaven is planned.

These standardized, as well as new tests to be developed specifically for Fairhaven or for colleges in general, will be employed as deemed potentially useful. As has been made clear by the rationale for strategies No. 1 and No. 2 above, a wide range of instruments is appropriate to strategy No. 3 which serves largely as a means to the ultimate goal of long-range behavior judged important by various groups.

After the student leaves Fairhaven College, the long-time longitudinal study will begin. Although it is perfectly reasonable to assume that any unique educational values of a small residential liberal college may have their main impact in less than four years, it is tempting to only consider those who finish a full four years at an institution. This does not seem a particularly sound procedure since, depending upon the maturity of the student involved, periods of varying lengths would be called for for different students. This study will include the students whose total tenure at the college is six quarters or more. Thus, if a particular student used a good part of one academic year before taking advantage of the educational freedom and then within another year moved far enough in his personal life that he no longer wished to remain in the college, this student would be included in the study. The periods of evaluation will be 5 years, 9 years, 14 years, and 29 years after the point of entrance in a class.

These studies of the individuals in later life would be guided by the same concern for breadth and for translatability into common language as has been emphasized above.

Remarks by Paul Woodring

In the last two paragraphs of his paper on evauation, Dean Charles Harwood mentioned briefly his plans for the long-range follow-up of Fairhaven graduates and former students. His plans were considerably more detailed than is indicated in this paper, and he had already made overtures to the Ford Foundation for support of such long-range evaluation. I have no doubt that other members of the Fairhaven staff will carry on with the plans, though the responsibility has not yet been specifically assigned.

In talking with me, Dr. Harwood had expressed a hope that we might make a determined effort to learn just how the college experience influences the later life and personality development of an individual and, hopefully, which part of the total college experience has the greatest influence. We discussed the possibility of establishing a permanent fund in order that the continuation of the project might be assured, regardless of changes in personnel.

The plans are to collect data regarding the more obvious facts–marriages, divorce, children, vocations, incomes, etc.–but to place the major emphasis on changes in personality and attitude which are admittedly much more difficult to assess. Procedural details have not been worked out, and the instruments to be used in assessment have not been agreed upon, but we would like to find answers to questions such as these:

1. Do students who are profoundly dissatisfied with the status quo translate their discontent into effective action when they move into positions of responsibility?
2. Do conventional students become conventional adults?
3. Does the usage of psychedelic drugs during the college years in any way affect later achievement or personality development?
4. Does unconventional sexual behavior during the college years have any effect on subsequent marital adjustment?
5. Are students who go through the experiences called "liberal education" more likely than others to take liberal views of social and political problems when they are middle aged?
6. Do "militant" students become militant adults?
7. What is the subsequent personality development of students who are highly resistant to all authority?
8. Do "square" students become more or less conventional as they grow older?
9. What traits, observable in college students, are predictive of serious personal or social maladjustment in later life?

The next steps will be to develop or select instruments for the identification of the various traits, characteristics, attitudes, or behavior patterns which are to be correlated with those to be assessed during the future years. Obviously, this will be a difficult task, but we are not ready to accept the view that it is an impossible one.

References

Educational Testing Service. *Institutional Functioning Inventory.* Princeton, N.J.: Author, 1968.

Educational Testing Service. *College and University Environment Scales.* (2nd ed.) Princeton, N.J.: Author, 1969.

Feldman, K. A., & Newcomb, T. M. *The impact of colleges on students.* San Francisco: Jossey-Bass, 1969. 2 vols.

Heist, P. A., McConnell, T. R., Webster, H. D., & Yonge, G. D. *Omnibus Personality Inventory.* New York: The Psychological Corporation, 1969.

Jencks, C., & Riesman, D. *The academic revolution.* New York: Doubleday, 1968.

Murray, H. A. *Explorations in personality.* New York: Oxford University Press, 1938.

Newcomb, T. M. et al. *Persistence and change: Bennington College and its students after twenty-five years.* New York: Wiley & Sons, 1967.

Sanford, N. (Ed.) *The American college.* New York: Wiley & Sons, 1962.

Stern, G. *College Characteristics Index.* Syracuse, N.Y.: Syracuse University.

White, R. W. *Lives in progress.* New York: Dryden Press, 1955.

Chapter 5

THE UNIVERSITY OF MICHIGAN'S RESIDENTIAL COLLEGE[1]

Theodore M. Newcomb
Donald R. Brown
James A. Kulik
David J. Reimer
William R. Revelle

Evaluations of experiments in education typically apply standard instruments designed to measure generally accepted criteria to student outputs such as academic performance and future educational attainment; often they simply accept volume of output against costs as the only true objective and practical measure. The specific instrumentation, methodology, and analysis may be quite sophisticated, but rarely are the criterion measures carefully chosen in terms other than by arbitrary considerations such as convenience, availability, and the possibility of normative data. A reading of Feldman and Newcomb (1969) gives ample evidence of the above in the area of student change studies and also makes clear the almost total absence of attempts at carefully chosen control groups within and across institutions. Jacob's (1957) conclusions regarding the absence of effect of courses in the social sciences were almost immediately seen as an example of the doubtful validity of inference drawn from arbitrary groupings of diverse studies using unspecified samples from a potpourri of widely differing settings (Brown, 1960a).

---

[1] All of the authors are associated with the University of Michigan's Center for Research on Learning and Teaching, whose facilities have made possible the work reported here. We are particularly grateful to Dr. Stanford C. Ericksen, the Director of the Center.

Certain studies have attempted to avoid these shortcomings—e.g., the Bennington studies (Newcomb, 1943; Newcomb et al., 1967), the Vassar studies (Sanford, 1956), and the Michigan State work (Dressel et al., 1959). The important feature of these studies is that attempts were made to state explicitly the nature of the educational setting or environment, the type of student input, and, most crucial, to specify the theoretical assumptions underlying the interaction among these variables in predicting observed outcomes.

In measuring outcomes of liberal arts educational settings where traditionally goals are accepted on faith and rarely stated explicitly enough to be operationalized, attention to these matters has a special importance. The liberal arts colleges have deep historical roots in their original development but, like all large institutions in a rapidly changing society, the often quite pragmatic origins evolve into a well-entrenched social reality which engenders emotional commitments and status functions quite beyond the original purposes. Indeed, the Mellon Studies of Vassar by Sanford and his associates (1956) were severely hampered by the inability of the researchers to elicit from the faculty of that institution agreement on the expected outcomes in their students as a result of four years at Vassar. Since the College was a well-established and smoothly functioning institution long before the researchers came on the scene, the evaluation could in no sense be said to be an integral part of the educational design. It was necessary to work backwards from intensive studies of alumnae on the one hand, and to relate independently measured attributes of students nominated by faculty as ideal students on the other, in order to specify the criterion variables to be evaluated (Brown, 1960b, 1956). Once these data were in hand, it was reassuring to discover that educational outcomes could indeed be operationalized along dimensions of student development, in ways that are congruent with psychological theories of personality growth and change, intellectual, social, and affective development in young adulthood.

Still, these evaluational studies were all after-the-fact impositions applied to an ongoing process. The program of evaluation described in this chapter differs in a major respect. It was planned from the very start of the three years of detailed discussions which led to the founding of the College. Two of us (Brown and Newcomb) were full members of the planning committee who were intimately involved in all decisions concerning curriculum, housing, governance, grading, and the College's relationships to the regents, administration, and faculty. The planners, moreover, were explicitly charged from the start to design a set of evaluational procedures which would not only provide a measure of attainment of the goals the committee set for itself and the new College, but would also enhance the experimental nature of the venture. The "experiment" was conceived of on two levels: first, to provide data for future assessment and perhaps modification of the experimental venture; and second, to develop a body of basic research data to add to the general pool of information now accumulating in the social psychology of higher education.

Thus, the evaluational design described here must be seen as it evolved out of the statement of purposes of the College (as described in the appendix at the end of this chapter). The evaluation proceeds on several levels in order to provide the most appropriate kinds of data at each level. These range from thorough descriptions of the entering students with appropriate control groups at the same university as well as at four other institutions (the latter are not here reported); ethnographic studies of the institutional "atmosphere"; analyses of attrition, course evaluation, student and faculty reactions to various aspects of the College; and detailed time budgets of student activity. In each instance the relevant data (the scope of which goes considerably beyond what is here reported) have been gathered with an eye to one or more of the purposes of the College and with careful attention to the appropriateness of methods and instrumentation. For those who are so fainthearted as to raise the "Hawthorne Effect" flag, we suggest careful attention to the control group design, cross-institutional comparisons, and the extensive use of normative instruments.

We now turn to the specific studies which tap the relevant levels of our enterprise. We shall begin with analyses of initial student responses to "standard" scales and questionnaires, and then turn to findings drawn from interviews and College records.

Characteristics of Entering Students

One cannot evaluate changes in students without knowing what they are like initially. An understanding of a college's "outputs" presupposes knowledge about its "inputs." And in comparing different colleges, or different parts of the same university, one must be informed about initial similarities and differences if one is to draw inferences about relative degrees and kinds of impact.

There are many grounds for assuming that students entering experimental residential colleges are distinct from those entering the larger, conventional units of our universities. Distinctions between these groups, rather than their commonalities, are emphasized here. Although commonalities of the two groups are many, the differences must be recognized, for these differences and their ramifications are important in the decision making of college administrators and research workers.

The First Entering Class

Since its establishment in 1967, the Residential College has provided students with an attractive alternative to education in a multiversity setting. Almost 20 percent of the applicants to the College of Literature, Science, and Arts (LSA) for the fall of 1967 indicated that they wished to be considered for admission as members of the first class of the University's Residential College (RC). For these applicants, the RC existed only as a description on a blurb—"a college within the college combining the environment of a small college with the resources of a large, cosmopolitan university." Fewer than one-quarter of the qualified applicants could be accommodated at the Residential College, however, and selection was carried out so that the students admitted to it were representative of the parent (LSA) student body in sex composition, in-state residence, and scholastic aptitude. In the comparisons that follow, a 91 percent sample from this population of 217 freshmen will be compared to a sample of students who chose to enter LSA. The two groups are similar in sex composition, residence, and scholastic aptitude.

At one point in the *College Student Questionnaire,* the entering classes at the Residential College and LSA were asked directly about their "personal philosophies" of higher education. They were asked to rank-order four philosophies, each described in a paragraph, according to the accuracy with which the philosophies described their own points of view. The philosophies were written to describe the norms of four college subcultures described by Clark and Trow (1966), and can be labeled Vocational, Academic, Collegiate, and Nonconformist.

In the analyses to be reported first, RC students are compared with students in two groups: those who had chosen the parent college, LSA, and a nationwide representative sample of students (from the CSQ manual). Table 1 presents the number and percentage of students choosing each type of orientation as nearest their "personal philosophy." The nationwide sample and LSA sample are fairly similar in their choices of educational philosophies. For both the nationwide and LSA samples, the most popular educational philosophy is the Collegiate. RC students differ from students in both of these samples in educational philosophy. They most often choose the

Academic philosophy to describe their educational orientation. They choose the Collegiate philosophy much less often than do students in the comparison groups, and they choose the Nonconformist philosophy at least twice as often. The differences between RC choices and LSA choices are highly significant by conventional statistical standards.

TABLE 1

Educational Orientations of Incoming Freshmen (percentages)

	Nationwide Sample	Residential College	College of Literature, Science, and Arts
Vocational	28	18.1	22.0
Academic	20	37.7	26.2
Collegiate	42	31.2	45.7
Nonconformist	4	13.1	6.1

The fundamental difference between students choosing an RC education and those choosing the conventional program at LSA can be more easily grasped if we consider a further comparison. Clark and Trow (1966) state that two of the educational types, Academics and Nonconformists, are concerned with ideas while the other two types, the Vocationals and Collegiates, seem less intellectually oriented. Both RC males and females are overrepresented in precisely those types which are described as concerned with ideas–the Academics and Nonconformists. Over one-half of the RC males and females choose the Academic or Nonconformist philosophy as best describing their personal philosophy of higher education, while fewer than one-third of the LSA students choose these orientations. This evidence thus indicates that RC students are more intellectually oriented than students preferring a more conventional program at the University of Michigan.

Comparisons were also made between the initial class of RC students and the LSA group on the psychometric scales of the *College Student Questionnaire* and the *Omnibus Personality Inventory*. Differences on 3 of the 5 CSQ scales and 5 of the 14 OPI scales reached significance at or beyond the .01 level. RC students scored significantly higher on Peer Independence, Liberalism, and Cultural Sophistication (all from CSQ), and in Thinking Introversion, Theoretical Orientation, Estheticism, Complexity, and Autonomy (from OPI scales). The scales that discriminate between groups are those that have common characteristics that may be labeled Anti-Intellectual Authoritarianism vs. Intellectual Openness. The mean scores on all differentiating scales indicate a higher degree of intellectual openness in RC students; i.e., lower on Practical Outlook and higher on all other scales.

The first comparisons–in terms of educational philosophy–indicated that RC students are more intellectually oriented than students choosing the conventional, large university environment for their education. The present analysis, based on a large number of personological measurements, supports these findings. Scales which indicate an interest in ideas, flexible use of intelligence, esthetic interests, and nonauthoritarian style are ones on which RC and LSA students differ. But entering RC students are neither better nor poorer in "adjustment" than LSA students; they are neither more nor less sociable.

In conducting these statistical analyses, we noticed that RC students were significantly (although not greatly) higher than LSA (control) students on SAT Verbal scores. This was unsettling since we knew that RC students and LSA students had been admitted so as to be equal in scholastic aptitude. The difference in SAT Verbal scores is apparently the result of "defections" from the Residential College and LSA after admission to these units but prior to matriculation by the students. For example, about 20 freshmen invited into the Honors Program and also into the Residential College chose the latter, whereas our LSA group contained none, to the best of our knowledge, who had been invited into Honors. Students in the LSA were more likely to join the Honors Program when invited than were students in RC. This in itself is interesting confirmation of the hypothesis that residential education captures the imagination of many of the more intellectually oriented among the University of Michigan's students. Nevertheless, the difference in SAT Verbal scores between RC and LSA students does not explain the personological differences in intellectual openness already reported. This was shown clearly in analyses which compared RC and LSA students on psychometric scales with verbal aptitude controlled. RC students were still significantly higher than LSA students on measures of intellectual openness when level of verbal ability was statistically controlled.

Four easily interpretable, large clusters of items emerged from a cluster analysis (Newcomb et al., 1970) that for our purposes had strong advantages over "raw scale scores." The scales based on these clusters have high reliability, are fairly orthogonal, and were labeled Anxiety, Intellectualism, Sociability, and Authoritarianism, on the basis of item content. The results presented in Table 2 show that RC students initially scored higher on Intellectualism than LSA students. That is, they are more likely to say that they like to read serious philosophical poetry; that they enjoy writing a critical discussion of a book or article; that they prefer to enter a profession which requires much original thought; that they like to make friends with persons who are rather sensitive and artistic; and so on. RC students are also lower on Authoritarianism, and are less likely to endorse items like the following: "God hears our prayers"; "No man of character would participate in sexual intercourse before marriage"; Every person ought to be a booster for his hometown"; "One of my main aims in life is to accomplish something that would make my mother proud of me." On Anxiety and Sociability, there are no differences between groups. The two groups are equally likely to report feelings of strain, tension, and going to pieces; and they are equally likely to report enjoyment of social gatherings, parties, and approaching other people.

TABLE 2

Educational Orientations of Residential College Freshmen in Four Years

	Percentage Entering Residential College in:			
	1967	*1968*	*1969*	*1970*
Vocational	18.1	10.6	7.8	7.3
Academic	37.7	29.2	26.9	29.3
Collegiate	31.2	32.6	29.9	25.0
Nonconformist	13.1	27.5	35.3	38.4

Characteristics of Later Classes

Data from one class at the University of Michigan are of limited value in making generalizations about the sort of students attracted to residential colleges in general. This is especially the case when the data from the residential college students are from the first class only; i.e., from students who chose a residential college for their education before the college began operation. For this reason, we looked at the second, third, and fourth entering classes at the Residential College. No control data are available for the second and third classes, but for the fourth group, entering in 1970, control data are available from students in LSA.

TABLE 3

Personal Characteristics of Incoming Freshmen

OPI Item Cluster	Residential College		College of Literature, Science, and Arts		Significance Level
	M	*SD*	*M*	*SD*	
Anxiety	51.7	17.9	50.5	17.4	n.s
Intellectualism	58.1	12.7	51.6	13.7	p<.001
Sociability	29.5	9.2	30.4	8.3	n.s
Authoritarianism	26.5	12.0	29.0	11.2	p<.05

Table 3 presents the numbers and percentages of entering students choosing Vocational, Academic, Collegiate, and Nonconformist philosophies of education in four entering classes at the Residential College. The trend is clear. There is a regular decline each year in the percentage of RC students choosing the Vocational and the Collegiate philosophies. While the Academic philosophy finds somewhat fewer adherents in the more recent Residential College entering classes, representation is still considerably greater than in a nationwide normative sample. The most striking trend in the choice of educational philosophies, however, is the movement of RC students toward a Nonconformist philosophy of education. Nearly 40 percent of the entering students at the Residential College in 1970 chose this orientation as nearest their educational philosophy–almost 10 times the proportion choosing this philosophy in a representative nationwide sample. In the same period, students entering LSA have not changed greatly in their choice of educational philosophies. For the LSA sample and for Honors students entering in 1970, the Collegiate is the most popular philosophy of education and the Nonconformist the least popular. Among students entering LSA in 1970, the proportions choosing educational philosophies concerned with ideas (Academic and Nonconformist) are: LSA, 34 percent; Honors students, 43 percent; Residential College students, 68 percent. The differences are highly significant by conventional standards.

The results from the psychometric instruments generally confirm the picture derived only from consideration of choices of educational philosophy among RC and LSA students. Comparison of the entering class of LSA in 1967 and the entering class in 1970 shows that few changes have occurred on the 5 scales of the CSQ and the 14 scales of the OPI. The LSA class entering in 1970 is significantly higher ($p<.01$) than the LSA entering class of 1967 on CSQ Liberalism and Social Conscience. The differences between entering classes at the Residential College in 1967 and 1970 are more striking. The RC entering class in 1970 is higher ($p<.01$) than the 1967 entering class on CSQ Family Independence, Liberalism, and Social Conscience, and OPI Complexity, Autonomy, and Religious Orientation[2]; this class is also lower on OPI Practical Outlook and Social Extroversion. Most of the changes can be interpreted as showing decreased authoritarianism in students entering the Residential College in the later year. The exception to this general rule is Social Extroversion, a variable that is not highly related to authoritarianism.

Students entering the Residential College in 1970, as compared with LSA entrants in the same year, scored significantly higher (at the .01 level) on Liberalism, Cultural Sophistication, Estheticism, Complexity, Autonomy, and significantly lower on Practical Outlook. These same differences were found for the 1967 entrants, but were accentuated and increased in 1970. Differences on Thinking Introversion and Theoretical Orientation are somewhat smaller than in the previous comparisons, but still reach the .05 level of significance. The difference on Peer Independence, previously reported, does not hold up for the class entering in 1970. Two new differences, not significant in the comparisons on the class entering in 1967, now appear. The RC students entering in 1970 are more independent of their families (CSQ Family Independence) and show less religious orthodoxy (OPI Religious Orientation) than the students entering LSA. Thus, in spite of such relatively minor differences, the comparisons reported for the original entering class at the Residential College in 1967 are highly similar to those found in 1970.

---

[2] A high score on this scale indicates low religious orthodoxy.

TABLE 4

Personal Characteristics of Four Classes of Incoming Freshmen at the Residential College

OPI Item Cluster	Mean for Classes Entering in:				*F*
	1967	*1968*	*1969*	*1970*	
Anxiety	51.66	53.55	56.06	56.75	3.3*
Intellectualism	58.11	61.59	62.33	59.85	4.9**
Sociability	29.49	27.70	28.07	25.69	5.3**
Authoritarianism	26.47	21.37	18.46	17.16	25.0***

*$p < .05$
**$p < .01$
***$p < .001$

The results from the four cluster-scores of the OPI presented in Table 4 provide a synopsis of a much larger set of results. The four entering classes at the Residential College do not differ greatly in anxiety, intellectuality, and sociability, although there is some trend toward increased anxiety and intellectuality and decreased sociability in successive entering classes. Authoritarianism, on the other hand, drops consistently and sharply in successive entering classes in the RC, and this difference is large enough to matter–it amounts to three-quarters of a standard deviation. The LSA entering classes of 1967 and 1970 also differ little in anxiety, sociability, and intellectualism, but the latter class is significantly lower on authoritarianism. This drop on authoritarianism is, however, much less than that for RC classes–between one-fourth and one-third of a standard deviation. The cluster scores thus give the following picture: the greater intellectualism found to characterize the initial entering class of the RC is maintained in successive classes at the Residential College; on authoritarianism, these classes drop sharply in successive years; and the difference in authoritarianism that characterized the 1967 groups at RC and at LSA is greatly accentuated by 1970.

Discussion of Results

All our analyses point to a general area of differences between residential and other students at the University of Michigan. They are intellective differences: differences in intellective interests and in intellective style. These differences exist over and above differences in intellectual ability, in which residential students may also be slightly higher. As important as these differences are, they should not obscure the equally important similarities between residential college students and their LSA peers. The two groups are similar in sociability and adjustment. The basis for choosing a residential college is neither affiliative need nor maladjustment; it is clearly related to the satisfaction of certain intellectual interests and needs.

Data presented by Heist and Bilorusky (1970) comparing students at four cluster colleges with students at the parent institutions in different parts of the country suggest that this phenomenon of self-selection is not a local event limited to the state of Michigan. Heist and Bilorusky's data–from Hofstra University, Nasson College, College of the Pacific, and the University of California–are remarkably consistent with our data. At the cluster colleges of these four institutions, students are consistently higher on OPI scales most indicative of intellectuality–Thinking Introversion, Estheticism, and Complexity. They are also consistently different on the "purer" measures of authoritarianism in the OPI–Autonomy, Religious Orientation, and Practical Outlook. In all cases, directions of differences are consistent with our findings, and the differences hold when analysis is done separately for the sexes. Furthermore, there are no other consistent differences in Heist and Bilorusky's data–on scales, for example, that tap adjustment and sociability.

Such data are of importance for both college administration and for research in higher education. College administrators are aware that the quality of an institution is determined in large part by the quality of the applicants to the college. They also know that rational curriculum planning demands that the curriculum be tailored in some sense to the needs of the incoming students. Our data indicate that residential colleges attract interesting and interested students. The promising students who enter residential colleges are "good bets" for further intellectual accomplishment and achievement. While this may make the public relations job of selling residential education easier–the students may shine on outcome measures simply on the basis of their original potential–it makes the design of an adequate curriculum especially challenging.

Preconceptions of the Residential College

For students applying to the University of Michigan for admission in fall 1967, the Residential College was no more than a description in a blurb. Some students were attracted to the college by the description, and other students chose not to become involved in this experiment in higher education. What expectations did the description of the RC generate in students? Students entering the RC and LSA in 1967 communicated their expectations to us before matriculation by responding negatively or positively to statements about this college environment they were about to enter (*College and University Environment Scales,* CUES). Such descriptions provide some information about the developing image of the RC, and they help explain why some of the most intellectually oriented students among the University of Michigan applicants chose to enter the RC, despite attractive alternatives.

We have reported elsewhere (Newcomb et al., 1970) the expectation that these students had of their college. Briefly, the commonalities in description of the two units in 1967 seemed more impressive than the differences, and the differences were described as "tentative but suggestive." Differences between the two groups on CUES scales, in fact, did not reach acceptable levels of significance in conventional group comparisons. There was some reason to believe that personality characteristics of the respondents were affecting their responses to CUES scales, and that the different personal characteristics of entering RC and LSA students should be taken into account in analyzing such data. When personality differences of the two groups were statistically controlled, some differences in descriptions of the RC and LSA could be found, and these differences reached moderate levels of statistical significance. It should be remembered, however, that the absolute size of such differences was small, and that in most respects RC and LSA students entering in 1967 anticipated similar college environments.

The small differences that did occur suggested that students entering RC expected congeniality and fairness in social and academic relations more than did their peers at LSA. Interestingly enough, RC and LSA students differed not at all on such CUES scales as Scholarship and Awareness–scales which in one way or another characterize environments with a high degree of intellectual commitment and academic emphasis. On such scales, no differences could be found between students entering the residential and conventional units of the University of Michigan. The results seemed in a sense paradoxical. Although RC students were clearly more intellectually oriented than those in LSA, they seemed to choose the RC not on the basis of its intellectual promise, but on the basis of its promise of psychological intimacy. The bright, intellectually oriented students entering the RC in 1967 apparently expected to find congeniality, clarity of expectation, and fairness at a residential college, and saw residential education as an alternative to the relatively impersonal, competitive, and confusing educational experience at a large university.

Table 5 shows that by the second year, the RC image has become far more definite. When compared to the RC or LSA class entering in 1967, the RC entering class of 1968 expects much more community, higher levels of campus morale, a higher quality of student-faculty relationships, and less practicality. In their high expectations of community and campus morale, the RC class entering in 1968 is especially distinctive–two-thirds of a standard deviation higher than the RC or LSA class entering in 1967. On the scales most indicative of intellectual interest and commitment–Scholarship and Awareness–there are again no differences. The differences reported here are found and are highly significant also when attempts are made to control for such initial personality characteristics as authoritarianism and intellectualism.

The image of the Residential College is thus crystallizing even after one year. Certain elements in the initial picture drop out; other initial notes are amplified. Thus, a sense of community and campus morale–a goal of the RC planners–has apparently developed, and the fact has traveled. Before a year has passed, high community and campus morale are clear features in the image of the RC. The students entering the RC in 1968 believe that faculty-student relationships are closer at the RC, and they also realize that the entrepreneurial style is not admired at the RC (CUES Practicality). The RC, as of 1968, still has not developed an image of strong intellectual commitment and awareness; students see RC and LSA as highly similar on this dimension.

TABLE 5

Expectations of the College Environment of Incoming Freshmen at the Residential College

	Mean for Class Entering in: *1967*	*1968*	*t-ratio*
Practicality	10.02	8.13	7.49***
Community	12.07	13.24	–3.57***
Awareness	17.74	17.91	–0.71
Propriety	6.90	6.55	1.19
Scholarship	17.43	17.20	0.86
Faculty-Student Relations	15.84	16.62	–2.29*
Campus Morale	8.04	9.22	–6.28***

*$p < .05$
***$p < .001$

By way of obtaining behavioral indices of actual rather than imaged environments, we analyzed "time budgets" submitted by more than 30 sophomores from each of the four populations we were studying. In most respects, these populations differed very little, but RC students reported less time spent in class attendance, in recreation, and in being alone. But they reported more time in study out of class, in bull sessions, in informal contact with faculty, and being in their residence halls (which contain classrooms and faculty offices). These findings are consistent with our analyses of CUES scores.

Overall, there is much in the developing image of the RC to satisfy the original planners of the college. A sense of community and campus morale has apparently developed in the college, and become part of the folklore of the college passed on to prospective applicants for the University of Michigan. The image is apparently used by applicants as one basis for deciding whether they will "fit" in the RC. The image of community and closeness apparently has greatest pull for low authoritarian students, who have entered RC in greater numbers each year. It may be a source of some disappointment to the RC planners to note that intellectual commitment or awareness are no more emphasized in the RC image than in that held for LSA.

Changes in Attitudes and Personality Characteristics

Any evaluation of a college should of course be related to its objectives, which can presumably be stated in terms of observed changes in students. To use one phraseology, we set out simply to do a better job of achieving the objectives of liberal education than is ordinarily possible, we believed, in large "conventional" undergraduate colleges. In another terminology, we hoped to demonstrate that those "desirable" kinds of changes that have repeatedly been observed in the past of American undergraduates had occurred more dependably and possibly in greater degree than in most other institutions.

By the same token, any evaluation of an innovative college should attempt to relate its findings to the specific, planned features of that college. In our case this meant taking particular note of students' peer relationships, as related to their observed changes. After all, the college had been planned with the intent to engage stimulation and support from peers in promoting certain kinds of changes. We also expected that more frequent and informal contacts with faculty would further facilitate such processes.

It happened that the present evaluators of this college were social psychologists who believed that they had certain bodies of theory to draw upon in designing their studies, and in interpreting them. Present understanding of developmental processes in late adolescents and young adults suggests that they are typically moving simultaneously toward higher integration and more precise differentiation of personality. Such processes, in a college setting, ideally include the following. (1) The freeing of impulse through the opportunity to learn and manipulate the symbols of human experience in imagination through contact with literature, philosophy, and the arts while not directly committing themselves. Here we do not want to underestimate the value of empathy in the overall educational process. (2) Enlightenment of conscience to the point where the individual believes in what he *ought* to do because he has arrived at a moral code by reasoned judgment. Contact on intimate terms with a diverse faculty and a diverse curriculum with some common core, training, and disciplined analysis of thought, and a tolerant but committed faculty and administration who make their own values explicit to the student are invaluable here for these developments. (3) Differentiation and integration of the ego such that the student increases his scope while becoming at the same time more of a unity. That is, perceptions and thought become more and more differentiated and personal responses more and more discriminating and interrelated.

The theoretical grounds for expecting potent peer-group influence upon undergraduates are not particularly obscure, though empirical demonstrations of such influence upon student change are comparatively few (Newcomb, 1943; Feldman & Newcomb, 1969). Briefly, the most relevant theoretical assumptions are these. (1) Potentially influential peer groups tend to arise out of frequent interaction, which in turn is facilitated by propinquity, and out of existing similarity of important interests and attitudes. (2) Such groups tend in fact to be most influential when they are relatively small, homogeneous (in certain but not all conceivable ways), and relatively isolated from counter-influences. To these it may be added that group members initially brought together on the basis of existing attitudes may reciprocally influence one another in changing those attitudes.

In part, at least, we had such considerations in mind in planning the Residential College. We believed that both degree and direction of students' changes would be affected by the arrangements described above, and we hoped to be able to demonstrate such effects. We now turn to the empirical data concerning change.

We have seen that the prospect of attending this residential college attracts intellectual and nonauthoritarian students and that such an environment is described in terms dramatically different from those applicable to the conventional liberal arts college. We will now consider the impact the Residential College has upon its students in effecting changes in attitudes and personality characteristics. We will discuss impact in relation to normal developmental trends of late adolescence as well as effects normally attributable to college attendance. We will show that RC students are more satisfied with their college experience than are comparable students at the LSA and are also more satisfied with their faculty and fellow students. We will also demonstrate that initial differences in the attitudes and personalities of RC and LSA students are *accentuated* so that the two populations become more distinct as they progress through college. Developmental trends discussed by Feldman and Newcomb (1969) will be shown to occur at both institutions but to be accelerated at the Residential College. Various dimensions of "openness to change" will be discussed and shown to have different effects on the RC and LSA populations. Finally, the impact of the RC upon its second and third classes will be shown to substantiate findings of the first year. We conclude that, while college environment is an important contributor to student development, the student's personality and its interaction with residential environment must also be considered.

We began our study with the expectation that, because of the more informal living and studying arrangements at their college, RC students would be more satisfied than LSA students with their faculty, administrators, and fellow students. We also predicted that the environment of the Residential College would be more conducive to student growth and development than that of the larger university. RC students, therefore, were expected to show more evidence of change than an LSA sample on those attitudes and personality characteristics which discriminate between upper and lower classmen.

Newcomb's theory of accentuation also led us to expect that initial personality differences in the two student populations under consideration would increase where those differences were relevant to the college experience. Thus, we expected that differences in intellectual and nonauthoritarian tendencies between the RC and LSA would be accentuated by the different environmental presses of the two colleges.

Finally, we felt that students with certain personality traits would be more receptive to change than others. Feldman and Newcomb's review (1969) suggested that more intellectually oriented students would change more on our various attitudinal measures than would less intellectually oriented students. In the same fashion, less authoritarian students might be more predisposed towards attitudinal change than more authoritarian students. Another "openness" variable suggested by Feldman and Newcomb which we considered was sociability. That is, those students who were more socially "open" might change more than those who were less socially "open."

The satisfaction of a college's students is one component of a school's success. Our data clearly show that seven months after the beginning of the academic year, students in the first class of the RC were much more satisfied than the LSA population with their faculty, students, and administration, as measured by the relevant CSQ scales. They also show higher scores on 29 of the 30 items making up these scales (22 of the items reach conventional significance levels). RC students are more likely than LSA students to say that their faculty is composed of superior teachers, who are genuinely interested in their students' personal and academic progress; they think that their administration fairly enforces the college regulations, which are logical and necessary; and they feel that their fellow students are honest and not overly involved in extremist or nonconformist activities.

Although students in the second class at the RC do not express such strong feelings about their satisfaction with the faculty and the administration of the RC, they are still very much higher than the LSA group from the first year (Table 6). The first and second classes do not differ significantly in the amount of Satisfaction with Students that they report.

TABLE 6

CSQ Satisfaction Scales for RC '71, RC '72, and LSA '71

	RC '71		RC '72		LSA	
Scale	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*
Satisfaction with Faculty	30.58	4.02	29.20	4.06	24.95	4.09
Satisfaction with Administration	30.74	3.87	29.23	4.68	28.04	4.19
Satisfaction with Students	31.33	5.00	29.28	3.67	27.84	4.68
N =	150		130		167	

Satisfaction is also related to initial predispositions. The more intellectual students express greater satisfaction with their peers than do the less intellectually oriented, but express less satisfaction with the administration. On the one hand, they do not feel that the student body is too nonconformist, extreme in political beliefs, or extreme in behavior; they feel that the student body is politically concerned, academically honest, and neither under- nor overcompetitive. On the other hand, these more intellectually involved students are more critical of the administration's rules and regulations.

A relationship is also found between Authoritarianism and the satisfaction scales. More authoritarian students are more satisfied with the administration as well as the faculty (LSA students only) of the university, but are less satisfied with their fellow students than are less authoritarian students.

The only relationship, for the second class, between initial predispositions and satisfaction that reaches statistical significance is the negative correlation between Satisfaction with the Administration and initial Intellectualism. This is similar to the findings for the RC class of 1971 as well as the LSA group.

In their review of studies of student development and the impact of colleges upon students, Feldman and Newcomb conclude that "No single principle has emerged in so many different guises as that of *accentuation*," which they see as an environmental reinforcement or extension of those characteristics of an individual that "selectively propel him toward particular educational settings." In its simplest form, this concept can be taken to predict that *any* differences between students in two academic environments will tend to be increased as the students progress through college. A more complex form of the hypothesis would state that those differences relevant to the academic experience will be accentuated. In either case, our data lend striking support to the hypothesis: of the 5 CSQ attitude scales, the 4 personality clusters from the OPI, and the 14 scales of the OPI, RC students change more in the expected direction than LSA students on 20 out of 23 comparisons. The three exceptions to the accentuation hypothesis are Peer Independence (CSQ), Altruism, and Response Bias (OPI). Students change as predicted in 4 out of 5 of the CSQ scales, 4 out of 4 clusters, and 12 out of 14 OPI scales (Table 7).

These trends continue for the second and third classes at the RC. Although there are no control groups for these two years, when compared to the LSA control group for the first class, the accentuation effect may still be observed. The second class (the class of 1972), relative to the LSA, shows accentuation on 3 of the 5 CSQ scales, all 4 of the OPI clusters, and 9 of the 14 OPI scales. The class of 1973, after one year, shows accentuation on three of the five CSQ scales, and has not yet taken the OPI.

TABLE 7

Raw Change on CSQ and OPI Scales

Scale	RC '71 *N*	RC '71 *Mean*	RC '71 *SD*	LSA '71 *N*	LSA '71 *Mean*	LSA '71 *SD*	*t*	Directions predicted by accentuation *RC LSA*
Family Independence (FI)	150	0.77	3.94	165	0.18	4.05	1.31	+
Peer Independence (PI)	149	−0.99	3.41	165	−0.25	3.65	−1.85	+
Liberalism (L)	143	2.37	3.35	158	1.33	3.80	2.51*	+
Social Conscience (SC)	147	0.39	3.69	163	0.35	4.39	0.08	+
Cultural Sophistication (CS)	150	1.10	3.26	167	0.77	3.01	0.93	+
Anxiety	78	7.01	13.56	123	2.62	14.74	2.12*	+
Intell	78	1.64	9.44	123	−0.42	9.38	1.52	+
Sociability	78	−2.99	6.47	123	−1.85	6.96	−1.16	–
Authoritarianism	78	−9.96	9.04	123	−7.77	7.13	−1.91	–
Thinking Introversion (TI)	78	.44	5.23	123	−0.35	4.93	1.08	+
Theoretical Orientation (TO)	78	.35	3.74	123	−0.50	4.43	1.41	+
Estheticism (Es)	78	.95	3.36	123	0.74	3.20	0.44	+

[Continued]

TABLE 7 *[Continued]*

TABLE 7

Raw Change on CSQ and OPI Scales

Scale	RC '71 *N*	RC '71 *Mean*	RC '71 *SD*	LSA '71 *N*	LSA '71 *Mean*	LSA '71 *SD*	*t*	Directions predicted by accentuation *RC LSA*
Complexity (Co)	78	2.83	4.72	123	0.94	4.55	2.83**	+
Autonomy (Au)	78	5.00	5.53	123	4.34	5.06	0.87	+
Religious Orientation (RO)	78	3.53	4.96	123	2.38	3.48	1.92	+
Social Extroversion (SE)	78	−2.13	5.70	123	−1.28	5.83	−1.02	−
Impulse Expression (IE)	78	4.97	6.53	123	4.02	7.17	0.95	+
Personal Integration (PI)	78	−1.56	8.30	123	0.96	8.63	−2.05*	−
Anxiety Level (AL)	78	−1.45	4.22	123	−0.57	4.13	−1.46	−
Altruism (AM)	78	−0.72	4.79	123	0.49	4.80	−1.74	+
Practical Outlook (PO)	78	−2.69	4.82	123	−2.24	3.71	−0.76	−
Masculinity-Femininity (MF)	78	−1.55	3.72	123	−0.80	4.32	−1.26	−
Response Bias (RB)	78	−1.47	3.86	123	−1.01	3.98	−0.82	+

*$p < .05$
**$p < .01$

An alternative explanation for the accentuation effect, one that does not hypothesize environmental reinforcements of initial differences, is that of "openness" or susceptibility to change. Thus, if two colleges admit students who differ on some attitudinal measure, it is also possible that they differ on some dimension of openness to change. As we have already seen, students at the RC are higher at entrance than LSA students on the OPI cluster measuring Intellectualism. We also know that, when Intellectualism is statistically controlled for, there are almost no other differences between the two entering classes. Therefore, students high on the Intellectualism cluster may be considered more susceptible to change (see below). It is necessary to demonstrate that the accentuation effect is not an artifact of this initial difference. Controlling for initial differences in Intellectualism for the first class does not yield results different from those already cited. We can now say that the observed accentuation effect is probably not due to initial differences on some openness variable, but is in fact due to the initial differences between the two colleges.

Possibly more important than the accentuation effect is the question of whether or not the Residential College aids students in their normal process of development. Feldman and Newcomb cite several developmental trends experienced by students in the college. First, in almost all longitudinal studies that they reviewed, they found that students are less authoritarian, more liberal in their political beliefs, and also show a greater degree of independence as seniors than they do as freshmen; these shifts typically include a decline in religious orthodoxy as they progress through college. A second basic trend is the one most expected: students have more intellectual interests as seniors than they do as freshmen; seniors are more concerned with abstract ideas and have a greater interest in cultural events than do freshmen. A third developmental trend is towards a self-acceptance of impulsivity: seniors are more able to express their impulses in action than are freshmen.

Examples of growth on several of these dimensions may be found in our control sample (Table 7). The most pronounced change was in the nonauthoritarian direction on the Authoritarianism cluster. The mean shift on this scale for the LSA was 7.8 items, or a change from a pretest mean of 30.8 to a posttest mean after 18 months of 23.0. This shift represents a change of .66 standard deviations on the cluster. Even after such a pronounced change of most of the subjects' scores, the test-retest correlation of .80 indicates a striking stability of the rank orders of individuals in authoritarianism. On the OPI scales, this change towards nonauthoritarianism is demonstrated by a shift of .68 standard deviations on the Autonomy scale and −.49 standard deviations on the Practical Outlook scale. The Liberalism scale from the CSQ shows a change of 1.33 items from 26.95 to 28.28 or a change of .25 standard deviations. While not as large as the shift on the

Authoritarianism dimension, the change is still significantly greater than 0, and in the direction predicted by Feldman and Newcomb. Yet another attitudinal change at the LSA that seems to conform with the trends outlined by Feldman and Newcomb is the shift of 2.4 items on the Religious Orientation scale of the OPI. The mean for LSA freshmen was 13.0 and, after 18 months, their mean had shifted in the less religious direction to 15.4. This change amounts to .41 standard deviations of the pretest. However, the shift of intellectual interests that is frequently found as students go through college was not found for the LSA. In fact, the LSA students went down .41 items or an insignificant .03 standard deviations over the period of 18 months after the pretest. Finally, LSA students increased on Impulse Expression as predicted by Feldman and Newcomb. The change amounted to 4.0 items from a pretest mean of 26.9 to a posttest mean of 31.0. This was equivalent to .47 standard deviations of the pretest scale.

While change along the above dimensions seems to follow a normal developmental process, at least three factors that are related to change can be identified. First, particular college environments may facilitate change along particular dimensions; second, some students may be more predisposed to change along certain dimensions than other students; and third, certain types of personality predispositions may interact with particular college environments to aid the process of growth. To assess the impact that the Residential College has on its students, it will be necessary to disentangle these three factors, as well as the problem of initial attitudinal differences between the RC and LSA sample.

This last problem can be alleviated through the use of *residual change* scores, which remove by regression the relationship between initial position on an attribute and change on that attribute. When this is done, the facilitating effect of the Residential College becomes apparent (Table 8).

Presented in Table 8 are *t*-tests of the differences in the residual change scores for the RC class of 1971 and the LSA control group. On the CSQ the most striking result is that the RC students changed more than the LSA group on the Liberalism scale. Of the four clusters, RC students changed significantly more than LSA students on the Anxiety and Authoritarianism clusters. Although not quite reaching conventional levels of significance, RC students also changed more than LSA students on the Intellectualism cluster. OPI scales that show significant differences between the two groups are: Theoretical Orientation, Complexity, Autonomy, Religious Orientation, and Personal Integration.

TABLE 8

Adjusted Change on CSQ and OPI Scales

	RC '71			LSA '71			
Scale	*N*	*Mean*	*SD*	*N*	*Mean*	*SD*	*t*
Family Independence (FI)	150	–.10	3.68	165	–.89	3.84	1.88
Peer Independence (PI)	149	–.36	2.95	165	.00	3.23	–1.04
Liberalism (L)	143	.48	3.06	158	–1.05	3.35	4.14**
Social Conscience (SC)	147	.22	3.37	163	–.42	3.73	1.60
Cultural Sophistication (CS)	150	.17	3.09	167	–.41	2.93	1.73
Anxiety	78	2.50	12.60	123	–2.17	13.85	2.42*
Intellectualism	78	1.06	9.36	123	–1.45	9.09	1.90
Sociability	78	–.51	6.21	123	.83	6.61	–1.43
Authoritarianism	78	–1.92	7.75	123	1.40	6.66	–3.24**
Thinking Introversion (TI)	78	.46	5.27	123	–.51	4.78	1.36
Theoretical Orientation (TO)	78	.67	3.69	123	–.52	4.16	2.08*
Estheticism (Es)	78	.24	3.07	123	–.26	3.01	1.18
Complexity (Co)	78	1.12	4.60	123	–1.11	4.33	3.49**
Autonomy (Au)	78	.45	4.60	123	–.98	4.67	2.13*
Religious Orientation (RO)	78	1.05	4.21	123	–.51	3.28	2.96**
Social Extroversion (SE)	78	–.46	5.47	123	.56	5.58	–1.28

[Continued]

TABLE 8 *[Continued]*

TABLE 8

Adjusted Change on CSQ and OPI Scales

		RC '71			LSA '71		
Scale	*N*	*Mean*	*SD*	*N*	*Mean*	*SD*	*t*
Impulse Expression (IE)	78	.87	6.06	123	−.46	6.70	1.44
Personal Integration (PI)	78	−1.28	7.60	123	1.45	7.94	−2.43*
Anxiety Level (AL)	78	−.37	4.01	123	.47	3.76	−1.52
Altruism (Am)	78	−.68	4.77	123	.13	4.28	−1.26
Practical Outlook (PO)	78	−.46	4.63	123	.39	3.68	−1.44
Masculinity-Femininity (MF)	78	−.36	3.56	123	.41	4.09	−1.39
Response Bias (RB)	78	−.25	3.51	123	.15	3.55	−0.81

*$p < .05$

**$p < .01$

While these results seem to indicate that the RC has a greater impact and accelerates basic developmental trends more than the LSA, the analysis is open to the same criticism leveled at the data on accentuation. Again controlling for initial differences in Intellectualism (Table 9), we see that the change in Liberalism from the CSQ, as well as changes on Anxiety, Intellectualism, and Authoritarianism from the OPI clusters, remain. Observed changes in Complexity and Religious Orientation from the OPI also continue to show a college effect, but the magnitude of the college effect for such variables as Autonomy, Theoretical Orientation, and Personal Integration is reduced. The most impressive implication of these findings is that of the developmental trends discussed by Feldman and Newcomb; three are aided by interaction effects with the environment of the Residential College.

TABLE 9

Adjusted Change on CSQ and OPI Scales with Intellectualism Controlled

Scale	Sum Squares	df	Mean Squares	*F*
Liberalism				
College	145.6	1	145.6	14.2**
Intellectualism	67.8	3	22.6	2.2
Interaction	6.7	3	2.2	0.2
Error	2,943.4	288	10.2	
Anxiety				
College	721.7	1	721.7	3.9*
Intellectualism	45.4	3	15.2	0.1
Interaction	184.8	3	61.6	0.3
Error	35,457.8	193	183.7	
Intellectualism				
College	565.5	1	565.5	6.8*
Intellectualism	136.1	3	45.4	0.5
Interaction	553.0	3	184.3	2.2
Error	15,964.0	193	82.7	
Authoritarianism				
College	424.0	1	424.0	8.4**
Intellectualism	47.9	3	16.0	0.3
Interaction	263.6	3	87.9	1.7
Error	9,753.5	193	50.5	
Theoretical Orientation				
College	51.1	1	51.1	3.2
Intellectualism	22.5	3	7.5	0.5
Interaction	93.8	3	31.3	2.0
Error	3,078.1	193	15.9	
Complexity				
College	293.9	1	293.9	15.0**
Intellectualism	33.2	3	11.1	0.6
Interaction	91.7	3	30.6	1.6
Error	3,789.3	193	19.6	

[Continued]

TABLE 9 *[Continued]*

TABLE 9

Adjusted Change on CSQ and OPI Scales with Intellectualism Controlled

Scale	Sum Squares	df	Mean Squares	*F*
Religious Orientation				
College	132.4	1	132.4	9.9**
Intellectualism	84.3	3	28.1	2.1
Interaction	39.3	3	13.1	1.0
Error	2,574.8	193	13.3	
Personal Integration				
College	202.7	1	202.7	3.3
Intellectualism	32.9	3	11.0	0.2
Interaction	198.2	3	66.1	1.1
Error	11,944.7	193	61.9	
Anxiety Level				
College	15.3	1	15.3	1.0
Intellectualism	14.9	3	5.0	0.3
Interaction	8.1	3	2.7	0.2
Error	2,931.9	193	15.2	

*$p < .05$

**$p < .01$

This finding is substantiated by an examination of the results for the second class at the RC. When comparing the different amounts of residual change on each of the CSQ or OPI scales for the two classes, it is striking that the two classes differ on only a few of the 23 scales. The only difference reaching significance was that the class of 1972 changed more than the first class on Family Independence ($p < .005$). Two differences that did not quite reach conventional significance levels were change on Religious Orientation and Impulse Expression. On both of these scales, the class of 1972 changed less than the first class.

While the two classes at the RC do not seem to differ on the amount of change, the second class did undergo changes that were significantly larger than those found for the LSA class of 1971. (There was no control group for the second year at the RC, so it is necessary to make any comparisons with the control group for the first class.) RC '72 students changed more than LSA '71 students on Family Independence and Liberalism (CSQ), and on Theoretical Orientation, Complexity, Autonomy and (negatively) Personal Integration (OPI).

RC '73 students continue to demonstrate these trends, for on the CSQ they differed significantly on residual changes in Family Independence, Liberalism, and Social Conscience when compared to the LSA class of 1971. Comparisons between different years and different schools are open to a great deal of criticism, but are reported as an indication of the continuity of the change process for the first three classes of the RC. When initial differences in Intellectualism are controlled for, the results are not quite as impressive but still similar to those of the first class. The class of 1972 changes significantly more than the LSA class of 1971 on the Family Independence and Liberalism scales (CSQ), as well as the Intellectualism cluster and the Autonomy scale (OPI), even when initial differences in Intellectualism are controlled for. The class of 1973 continues this shift in the more liberal, nonauthoritarian direction by changing more than the LSA class of 1971 on the Liberalism scale of the CSQ.

We have seen that initial differences between the Residential College and the LSA are accentuated and that the RC seems to accelerate the process of normal student development. We turn now to the question of individual differences in susceptibility, or "openness" to change. Feldman and Newcomb cite six variables which may be considered as measures of openness towards change. Three of these dimensions are of special interest to us since they are measured by the OPI clusters. According to Feldman and Newcomb, students who are more *intellectually oriented* upon their arrival at college tend to change more on various attitudinal measures than do less intellectually oriented students. In the same fashion, less *authoritarian* students seem more predisposed towards attitudinal change than do more authoritarian students. Feldman and Newcomb also report that students who are more *socially open,* or sociable, tend to experience a greater amount of change than do less sociable students.

For changes on the CSQ after one year for all subjects from the LSA and from the first three years of the RC, two of our four personality clusters seem to qualify as measures of this openness. The initial level of Intellectualism correlates significantly and negatively with residual change on three of the five scales: Family Independence, Peer Independence, and Liberalism. It must be pointed out, however, that while these correlations are statistically significant, only 2 to 4 percent of the variance on the change measures is being accounted for.

Openness to change on the OPI for the LSA and the first two classes of the RC is related to initial Intellectualism, non-Authoritarianism, and Sociability. When all students are considered together, the pattern is one of intellectual students becoming less authoritarian, and of the less authoritarian students becoming more intellectual. Sociable students tend to shift in the nonauthoritarian direction on the Authoritarianism cluster. Another variable related to change on the OPI is sex. Females tend to change more towards a nonauthoritarian position than do males, as well as to change their religious beliefs to a less orthodox position.

In an earlier report of some of these findings (Newcomb et al., 1970), it was suggested that there are different variables related to openness for the LSA and for the RC. At the LSA, initial Intellectualism predicted change towards nonauthoritarianism, and initial Authoritarianism related negatively with change towards more intellectual interests. At the RC, however, the best predictor of change on the OPI is initial level on the Sociability cluster. RC students who are more sociable become much less authoritarian than do less sociable students. This effect becomes much stronger for the second class of the RC: about 25 percent of the variance on the Authoritarianism scale is accounted for by initial Sociability.

On the CSQ, however, the relationship of Sociability to change is not quite as clear. While for the first class, more Sociable students at the RC change significantly more on the Peer Independence and Liberalism scales, change on neither of these scales relates to sociability for either the second or the third class. For the three classes taken together, Intellectualism relates significantly to all except change on Liberalism, Authoritarianism relates significantly and negatively with both change on Family Independence and on Peer Independence, and Sociability relates to none of these significantly.

The appeal of the findings relating Sociability to change for the first class and the changes on the OPI for the second class is that, even though they are somewhat tenuous, they make sense psychologically. We have seen that students at the RC are initially less Authoritarian and more Intellectual than are students at the LSA. It seems from the first class's CSQ and OPI data, and especially the OPI from the second class, that those students who are more sociable and interact more with their fellow students at the RC are the ones who are most affected by it. Thus, the impact of the RC is greatest for those students who interact with it the most.[3] For the LSA, however, where social interaction is not necessarily related to classroom behavior, those students who are most intellectually oriented or hold less authoritarian beliefs are also the ones who are most open to the experience and impact of college. This hypothesis is more suggested by our data than confirmed by it and needs confirmation from other institutions that also have experimental colleges that attract students similar to those attending the Residential College.

Student Attrition

The capacity of an educational institution to retain students who freely choose to study within it is *ceteris paribus* certainly one measure of its success. Thus far, attrition in the University of Michigan's Residential College is, I suspect, somewhat higher than those who favor the creation of such environments would expect. Among the first two classes entering the Residential College, 25 percent of the first class and 30 percent of the second class had dropped out by the spring of the sophomore year. This slight increase in the attrition rate is somewhat surprising, since entering students in the initial class were presumably less well informed than subsequent cohorts about the nature of this new and distinctive setting. Thus, the probability of a mismatch between individual and institution was greater in the case of the first class. For example, the proportion in this group planning a major in the natural sciences who later dropped out of the RC is considerably higher relative to the second cohort, apparently because they were less aware that the Residential College curriculum stresses the social sciences and the humanities much more than the natural sciences. However, there is considerable evidence that unrealistic expectations continue to be an important cause of attrition.

The student who drops out of the Residential College has two alternatives: he may transfer to another unit within the University, or he can leave the University altogether. Approximately 56 percent of all dropouts have remained at Michigan, transferring in most cases to the College of Literature, Science, and the Arts. We estimate that after two and one-half years, about 20 percent of those who had entered in the first two classes had left the University altogether, which is lower than the LSA drop-out rate over the same period. Of those students who simultaneously drop out of the

---

[3] For a discussion of the relationship between Sociability on the OPI and behavior at the RC, see Newcomb et al., 1970.

Residential College and the University, it is estimated that at least 70 percent enroll in another school within one year.

Students of attrition have generally found no difference between the drop-out rates for male and female students, and this has also been observed at Michigan (in the College of Literature, Science, and the Arts). But the pattern of attrition in the RC is atypical in this respect. In both classes, the attrition rates for females have been much higher, about 33 percent after almost two years, compared to 20 percent for males. The apparent causes of attrition for the two sexes, however, are not very different.

In order to assess reasons for dropping out, a questionnaire composed of both closed and open-ended questions was devised and sent to most of the dropouts from the first cohort (although returned by a disappointingly small proportion of them), and an almost complete survey of dropouts from the second class was carried out (by Victor Omelczenko, a student in this class). In the case of most dropouts, some information bearing on their decision to leave is also available from RC administrative personnel and from the student's file.

The reasons for attrition are, not surprisingly, highly varied, and thus difficult to categorize. We have therefore classified dropouts only according to a very gross three-fold typology. About 70 percent of all students no longer in the College (on whom information is available) evidently dropped out either because they were dissatisfied or had a problem with academic life in the RC, or because they found the social environment incompatible. All that can be said about most of the other dropouts is that they left because of personal reasons—e.g., marriage (or pregnancy), financial difficulties, and psychological problems.

Reasons for attrition related to academic aspects of the Residential College are mentioned more frequently and rated higher in importance than reasons of any other type. The most common dissatisfactions or problems in this area are (1) that the curriculum is too rigid and structured—i.e., too many required courses; (2) that particular courses are not well organized, lack structure or direction, or are poorly taught; and (3) dislike for pass/fail grading or the use of written evaluations as contrasted with letter grades. There is little doubt that many students transfer to LSA because of the more varied course offerings and greater freedom to elect courses there—even though in many cases they could elect those courses as RC students. A number of students say they dropped out primarily because the major they had in mind (a science or preprofessional specialty in most cases) was not compatible with the RC curriculum. For some, the more stringent language requirement in the RC was apparently a major factor in their dropping out.

The social atmosphere which has evolved in the Residential College is unconventional, often portrayed by terms such as free-swinging and sometimes as "T-groupy." Some dropouts are inclined to describe it in much stronger language—one of them who says he enjoyed the RC socially and left after two years only because of his premed major remarked, ". . . as many RC people admit, it is becoming a 'zoo' of ultra-liberal, unconvinced, plastic wanderers." Another student who dropped out commented, "The people in RC are very strange, not all, but maybe 30-40 percent. It is not a cross-section of all walks of life, but more of the hippie types." It seems clear that many who have left the College did so because they preferred a more traditional social environment—a more private and stable atmosphere, less characterized by confrontation and nonconformity. Females especially indicate that they desired more anonymity than was available in the RC community. Some dropouts feel that the RC is "too narrow a community" and is too cut off from the larger university. "Too small a group! I get tired of having the same people in all my classes," one dropout remarks.

Let us now turn to relationships found to exist between attrition and various psychological indices. Items and scales from the following inventories were analyzed: the *College Student Questionnaire* (CSQ), the *College and University Environment Scales* (CUES), the *Omnibus Personality Inventory* (OPI), and the *Opinion, Attitude and Interest Survey* (OAIS). Measures at time of entrance were provided by all four inventories; the CSQ and CUES were also taken again at the end of the freshman year. If all students who dropped out of the Residential College are lumped together and compared with all students still in the College, almost no significant differences are discerned. However, males who leave the RC differ from males who remain and females are also dissimilar. Moreover, within the dropout population itself there exist contrasting subgroups, especially when they are dichotomized according to destination.

Students who leave the Residential College but remain in the University are, irrespective of sex, sharply distinguished from students who drop out of the University altogether. These two groups differ at time of entrance: students who remain in the University are initially more authoritarian, more concerned about getting good grades, and more inclined to espouse a collegiate educational philosophy. The dropouts who leave the University altogether are initially much more independent of both family and peers and are more intellectual; in their educational philosophy they are more committed to nonconformity, individualism, and self-discovery. However, with respect to reasons given for dropping out, these groups are only slightly dissimilar. Problems of social incompatibility are more evident in the case of those students who did not leave the University; pass/fail grading in the RC was also a more important reason for this group.

These two kinds of dropouts differ from students still in the RC, but in different ways for males and females. Girls who dropped out of the University altogether, as compared with girls still in the College, are considerably more independent of both family and peers, more willing to express their impulses, and less concerned about having smooth interpersonal relations or being popular; they are also somewhat less authoritarian and more intellectual, and are more inclined toward a philosophy of nonconformity. On the other hand, girls who transferred from the RC to another unit in the University were upon entrance quite similar to girls still in the RC, except that the former expressed a much greater interest in the pursuit of vocational goals. But on measures of environmental perception and satisfaction obtained at the end of the freshman year, a number of highly significant differences are found. Females who have remained in the RC describe its environment as being more friendly and supportive, and characterized by group cohesiveness; they are more inclined to describe the RC atmosphere as scholarly, as one in which faculty set high standards but are also flexible and helpful. Girls still in the College were also more satisfied with the faculty and with the administration at the end of their freshman year than were girls who subsequently transferred to another unit within the University.

Male students reveal a somewhat different pattern. Males still in the RC and males who left the University are very similar on both initial and subsequent measures. The latter group was higher initially on family independence, but this is the only significant difference. However, when dropouts who stayed at Michigan are compared with students still in the RC, the most pronounced differences between dropouts and nondropouts emerge. Males still in the RC are considerably less authoritarian and more intellectual, they initially achieved higher scores on both the verbal and math portions of the *Scholastic Aptitude Test* (SAT); they also come from a higher socioeconomic background than the dropouts, and exhibit a higher level of cultural sophistication.

Academic failure has often been cited as a leading cause of attrition, and in one study (Suczek & Alfert, 1966), it was found that the personality profile of dropouts with failing grades was quite different from that of dropouts in good standing; the main difference was that failing students were more authoritarian. In the present study, however, academic failure is neither a common cause of dropping out, nor are dropouts with poor academic records particularly different from other students. Only about 15 percent of all dropouts were forced to leave the Residential College for this reason or were on academic probation at the time they dropped out. (Most students who have been suspended for academic reasons have later returned to the RC, and are not classified as dropouts.) The "academic failures" in our drop-out sample tend to be from a lower socioeconomic background and to have less satisfactory study habits than other dropouts, but their scores on the SAT are not significantly lower than the latter group.

As noted earlier in this paper, variables associated with both self-selection and "change" variables include authoritarianism, independence, intellectualism, and impulse expression. These same variables also tend to distinguish students who remain from students who drop out of the RC, and also to discriminate within the drop-out population. Perhaps the individuals who later find in the Residential College a compatible academic and social environment are, in terms of their psychological makeup and stage of development, a somewhat restricted group, even within the select population of potential Michigan students. Our findings suggest that students coming into the RC who are outside of a given range on several dimensions–independence, authoritarianism, intellectual orientation, and impulsivity (primarily)–are less likely to remain in this setting. Apparently, individuals who are either more or less mature or "developed" than the average RC student in terms of these variables do not get along as well in this environment.

A finding analogous to this one was noted by Gurin, Newcomb, and Cope (1968) in the College of Literature, Science, and the Arts at the University of Michigan; i.e., that female students of two contrasting types were more likely to drop out, women from less cosmopolitan backgrounds having a low intellectual-aesthetic orientation and also women of high cosmopolitan backgrounds who were more intellectually oriented than the average. They also noted similar, but less significant tendencies, with respect to religious and political orientations in interaction with background. But our data also show that many dropouts are quite similar to students who remain in the RC, and further research is clearly needed, both to determine whether the attrition of these students is perhaps related to other personality variables (not explored in this study) and also to validate the reliability of relationships suggested above.

Assessment of the Curriculum

The College's "Core Curriculum" was a special area of evaluation in 1969-70. The initial set of requirements was designed by a Planning Committee, composed of a dozen members of the faculty of the parent college (Literature, Science, and the Arts), beginning in 1964–three years before the actual opening of the Residential College. The College's general mandate was to work toward the general objectives of the parent college, but to achieve those objectives "in our own way," taking advantage of the opportunities offered by the creation of a new, innovative enterprise.

By the end of the second year, students' dissatisfactions with the required "core" were becoming highly audible. One of the special targets was the language requirement, which was somewhat more demanding than that of the parent college. Another was "Language and Logic," required of all freshmen, meeting as a single body with no small sections, and devoted mainly to lectures. It was also considered extremely demanding, in terms of amount and difficulty of required reading, and there was resentment because, atypically, students received letter grades (instead of pass/fail), based exclusively on written examinations. (This atypical format had been decided upon because its low cost per student would help to offset the high costs of the small Freshman Seminars, also required.) "Human Behavior," our course designed to meet social science distribution requirements, and "Western Man," serving to meet the same requirements in humanities, each met both for long lectures and for small discussion groups; these were widely regarded as simply not good courses, and there were many complaints because the requirement permitted no alternatives. Thus, the only feature of the Core Curriculum that was generally considered satisfactory was the Freshman Seminar.

After several such complaints had been aired in the Representative Assembly (the College's governing body, half of whose members were students), it was decided to create a Core Curriculum Review Committee. This body, made up of approximately equal numbers of students and faculty, spent some six months in reviewing objectives, in informal soundings of community opinion, in preparing, criticizing, and revising proposals submitted by members of the Committee as well as by others. Eventually, it unanimously approved a formal report which, in principle, was adopted by the Representative Assembly. It was far-reaching, and in some respects "radical." It retained the notion of a "core curriculum," but introduced much greater flexibility, both in providing alternative paths toward meeting every requirement, and in permitting more latitude as to times when requirements should be met. The more extreme provisions in the report were that the core program should be recommended rather than required. These provisions cannot be formally adopted without further conversations with the parent college.

So far, we have described processes of curricular change based on very informal kinds of evaluation, but we also arranged systematic interviews with stratified samples of students (100 of them, representing nearly a fifth of the then population), and of faculty (half of 80-odd of them). These interviews were carried on by graduate students from the University's Center for the Study of Higher Education.

Summaries of long reports on student and faculty interviews are appended hereto. In many ways, the findings parallel the expectations of the Review Committee, but in other respects go beyond them; e.g., the conclusion that students' concern with the development of identity, humanity, and meaningful action on social and political issues is more prominent than narrowly defined academic concerns. Or that students' interaction with each other, and to a somewhat lesser extent with faculty, is the most central influence in their college experience—more often than inside the classroom. Or that "While courses are on occasion important experiences, they clearly are not central in the students' experience."

Such findings are, in a way, not surprising, since the College has been designed to maximize interaction among its members. But they are also sobering, because they were not fully foreseen by many community members. We had planned a college in which free and informal interaction would occur, in academic as well as in other ways. Had our planning somehow gone awry?

An outright majority of the faculty felt that the objectives of the Core Curriculum had not been met very well, particularly insofar as those objectives included common knowledge expected of "any liberally educated person." Many of them attributed this relative failure to the rigidity of the curriculum, together with students' antipathy engendered by what they considered arbitrary requirements. Several of them agreed with the considerable number of students who felt that the required core—established in advance by a Planning Committee of faculty from the parent college—was incompatible with the College's policy of self-government shared by students, faculty, and administrators. Nevertheless, a general sense of uneasiness about academic standards was fairly common within the faculty.

Both students and faculty were queried about student participation in decisions about the conduct of courses in which they are enrolled. Typically, a teacher and his students in a given class were in close agreement about relative influence in classroom decision making–and, given such congruence, students are generally satisfied with the course. If, however, students and teacher differed about this, most students preferred the traditional teacher-dominant role. And, interestingly enough, when both teacher and student perceived that students were dominant in making decisions, student satisfaction was markedly low.

Summary of Student Interviews[4]

The central curricular issue for the Residential College from the students' viewpoint is the extent to which the college's core curriculum should be required of all students in the college. While there is clear support among almost all of the students for retaining a clearly designated core curriculum for the college, there is broad support for redefining the nature of this curriculum. Rather than having it consist of a set of prescribed courses with few if any alternatives, most students would prefer to have at least a selection of courses from which they can choose within designated areas of study if not simply a recommended curriculum with no set requirements at all.

While satisfaction with their Residential College experience in general is relatively high, satisfaction among students with the core curriculum is strikingly low. The source of this dissatisfaction seems not to rest with the basic rationale of the core curriculum or with the faculty available to present it but rather with the specific implementation of the concept in a set of required courses. The required nature of the curriculum together with the experienced inadequacy of many of the courses themselves has led to student discontent and disenchantment with this aspect of their Residential College experience.

The issue of flexibility vs. structure seems to be a central one in the academic endeavor of the college. The Residential College as a whole is characterized by a great deal of student freedom and responsibility and this finds expression in the individual courses as well, where students are often able to work independently in areas of their special interest. Yet the required nature of the curriculum is often viewed by students as inconsistent with the entire spirit of the college and thus becomes a point of contention. Moreover, the presence of so many students within a course who have not been attracted to it due to its specific content or focus seems to have led in many instances to an attempt to accommodate the varied interests within the group to the point where the course itself lost all sense of direction and integrity. This situation is reflected by students who decry the lack of structure in courses and express a sense of frustration with this state of affairs.

---

[4] Excerpted from a summary of a longer report by William Moore, of the University of Michigan's Center for the Study of Higher Education, who not only had primary responsibility for preparing the interview schedule, but also took general responsibility for organizing the staff of interviewers from that Center, with help from Professor Robert Blackburn.

The students' characterizations of their collegiate experience are noteworthy in that they tend not to reflect the more traditional academic interests and goals but rather stress a more general concern with personality growth and development. A concern with interpersonal relations and social and emotional development proved as important as development in the purely intellectual sphere. The community of the Residential College then seems to be not so much an intellectual one as a more generally personal and political one, in the sense that concern with the development of one's identity and humanity and concern with meaningful action on social and political issues is more prominent than narrowly defined academic concerns. Furthermore, it appears that the students' contact and interaction with other people in the college–primarily students, but to a significant degree faculty as well–is the most central influence in their collegiate experience, with this influential more often outside the classroom than inside. While courses are on occasion important experiences, they clearly are not central in the students' experience.

Summary of Faculty Interviews[5]

The majority of the faculty are happy with their College experience, finding it interesting, absorbing, and fulfilling. A majority of them report that they have devoted more time to the College than contractually expected. Most important to them are personal contacts with students, and the chance to experiment with innovations in teaching. In spite of certain dissatisfactions and misgivings, they are reasonably content with the College's condition and direction.

---

[5] Summarized and in part excerpted from the report by Dr. J. Bruce Francis, now of the Department of Higher Education at SUNY (Buffalo, New York), who constructed the interview schedule and carried out all of the interviews.

Seventy percent of those interviewed felt that some sort of core curriculum is essential, but only 15 percent were satisfied with the existing core. Most faculty members accepted both a "common knowledge" objective and a "personal development" objective–but with somewhat different emphases, and even different understandings of the terms. The differences had to do mainly with relative emphasis upon "What every liberally educated person should know" and upon "The development of cognitive skills and of psychological strength to deal with complexity," regardless of content. More than half of the respondents thought that foreign language study should no longer be required, and minimized the importance of "stable core courses providing sequentiality" and "known foundations, without gaps or duplications."

Faculty members who emphasized "traditional liberal arts goals" attributed the deficiencies of the core to "poor execution of courses," whereas those who stressed "free discussion, both in and out of class leading to growth of academic community" saw the core's failures as "stemming from its rigidity, the rebellious attitudes engendered by its being required, and the persistence of faculty-centered, traditional motives of education." In spite of such differences, 80 percent agreed that "The Core Curriculum should be more flexible in terms of course options, sequences, and time required for completion"; only 12 percent disagreed.

Asked about the College's students, many respondents noted two types. The first consists of a rather small proportion of outstandingly open, intelligent students who thrive on independent work. Such students are believed to be more common in the RC than in its parent college (most faculty members teach or have taught in both places), and some feel that such students' brilliant work is facilitated by conditions at the RC. A second type is "the highly vocal, articulate, and often radical student who involves himself deeply in political activities within the College, and who often is openly critical of traditional, 'Establishment' norms." They are highly visible, and often exert a good deal of influence. Some faculty, who are apt to regard such students as arrogant and intolerant, report the prominence of this type.

Third and fourth types, less commonly noted, consist of those who are "average" and not very conspicuous, and, fourth, of a small group who are "lost or drifting." There appear to be special conditions that facilitate the development of all but the third type in this College. It is also worth noting that many teachers frankly reported their own initial discomfort, and subsequent problems of adaptation, in a college characterized by student informality and "lack of distance" between teacher and student. The typical response, as reported by teachers themselves, has been a good deal of change on their own part.

References

Brown, D. R. Some educational patterns. *The Journal of Social Issues,* **12**(4), 1956.

Brown, D. R. Social changes and the college student. *The Educational Record,* **41**(4), 1960. (a)

Brown, D. R. Non-intellective qualities and the perception of the ideal student by college faculty. *Journal of Educational Sociology,* February 1960. (b)

Clark, B. R., & Trow, M. The organizational context. In T. M. Newcomb, & E. K. Wilson (Eds.), *College peer groups: Problems and prospects for research.* Chicago: Aldine, 1966.

Dressel, P. L., Lehmann, I. J., & Ikenberry, S. O. *Critical thinking attitudes and values in higher education.* East Lansing, Mich.: Michigan State University, 1959.

Feldman, K. A., & Newcomb, T. M. *The impact of college on students.* San Francisco: Jossey-Bass, 1969.

Gurin, G., Newcomb, T. M., & Cope, R. G. *Characteristics of entering freshmen related to attrition in the literary college of a large state university.* Office of Education, U.S. Department of Health, Education, and Welfare Project No. 1938. Ann Arbor, Mich.: Survey Research Center, Institute for Social Research, University of Michigan, 1968.

Heist, P., & Bilorusky, J. A special breed of student. In J. G. Gaff (Ed.), *The cluster college.* San Francisco: Jossey-Bass, 1970.

Jacob, P. E. *Changing values in college: An exploratory study of the impact of college teaching.* New York: Harper and Brothers, 1957.

Newcomb, T. M. *Personality and social change: Attitude formation in a student community.* New York: Holt, 1943.

Newcomb, T. M., Brown, D. R., Kulik, J. A., Reimer, D. J., & Revelle, W. R. Self-selection and change. In J. G. Gaff (Ed.), *The cluster college.* San Francisco: Jossey-Bass, 1970.

Newcomb, T. M., et al. *Persistence and change: Bennington College and its students after twenty-five years.* New York: Wiley and Sons, 1967.

Sanford, N. Personality development during the college years. *The Journal of Social Issues,* 12(4), 1956.

Suczek, R. F., & Alfert, E. *Personality characteristics of college dropouts.* Washington, D.C.: Educational Research Information Center, 1966.

APPENDIX

Brief Description of the Residential College[6]

The Residential College, established in 1967 within the College of Literature, Science, and the Arts, offers a four-year program in the liberal arts animated by a spirit searching out new and better modes of education, a spirit grounded in the conviction that this search is most fruitful when it is the common endeavor of the whole community, and when it encourages the personal growth of all. Its community is the heart of the Residential College.

Governance. The Residential College has neither an exclusive faculty-administration government nor an exclusive student government. Rather, managing the College through its Representative Assembly and committees responsible to it, the several estates of the community–students, faculty, administration–are effectively represented in all these bodies.

Course of study. Students in the Residential College normally enter a course of study centered on a core curriculum, supplemented by elective courses in the Residential College and the College of Literature, Science, and the Arts, and culminating in a concentration program chosen from among those available in the College and its parent unit.

The core curriculum is intended both to provide exposure to diverse areas and skills and to bring students in touch with ideas and ways of thinking likely to prove of personal value now and in the future. Both the structure of the core curriculum and its specific courses are being reviewed in the light of the experience of the past three years.

Students are expected during their academic career to elect appropriate courses in humanities, social sciences, and the natural sciences to fulfill the distribution requirements required by LSA. It is the policy of the Residential College that independent study of some kind is an educationally desirable experience for all students.

All students must at some time elect a *practicum,* or its equivalent, in the creative and performing arts.

Each student must achieve reasonable fluency in one foreign language. After a student attains fourth-term proficiency, he is expected to take an additional course conducted in a foreign language in the field of his choice.

Students are graded "pass" or "fail" in most courses. Instructors in such courses provide students with detailed written evaluations of their performance, and these evaluations become an integral and permanent part of a student's official academic record.

---

[6] Excerpted from an official leaflet, prepared in 1970.

Concentration. Concentration programs (majors) are available in most of the areas represented by the departments of the College of Literature, Science, and the Arts. Individualized interdisciplinary concentration programs can be arranged with members of the faculty. The Residential College is developing new concentrations, not now available in the parent unit.

Faculty. Instruction in the Residential College is given by faculty who are also members of departments of the College of Literature, Science, and the Arts, and who have a special interest in undergraduate and innovative education. These teachers at times work closely with students, both formally and informally.

Degree. On completion of the Residential College's requirements for graduation, students will be awarded the Bachelor of Arts degree by the College of Literature, Science, and the Arts. Students who complete a concentration in science or mathematics may be eligible for the Bachelor of Science degree.

Living. The Residential College occupies all of the completely remodeled East Quadrangle, one block south of the central campus, in which classrooms, offices, and language laboratory facilities were constructed. The College expects eventually to grow to 1,000 students. All freshmen and sophomores must live in the Quadrangle; juniors and seniors may do so at their option.

Admission. The freshman class is carefully selected to create a cross-section of the student body of the College of Literature, Science, and the Arts. Anyone seeking admission as a freshman in the College of Literature, Science, and the Arts may at the same time designate the Residential College as the unit in which he wishes to work toward a degree.

Chapter 6

EVALUATING CHANGE IN NEW COLLEGES

D. Gordon Rohman

Although the group of colleges represented in this anthology has been variously described as residential, cluster, inner, experimental, or experimenting, I like the simple designation "new."

All the colleges here are new, each having been created within the last decade or so. Apart from being merely descriptively true, however, "newness" offers a useful clue to those, such as evaluators, who seek to examine our style of life.

The first thing "being new" means is that we have all been given a rare and precious opportunity: to make a deliberate and conscious choice of an alternative future. Not many men in their lifetimes get such a chance. Not many organizations either. And even though the scope of our actual choices may have ranged from the narrowest set of options to a blank check, we nevertheless all got some real chance to give a new answer to an old question: what is the best way to school it?

Each of our colleges is our answer. Probably not our ideal answer, but a possible one for us in our time and place. Some of us stayed pretty close to shore, others waded out—some perhaps over their heads! But in varying amounts of wetness, we all got into some aspect of the change game. What interests me as much as the outcome of that game is what we have learned about playing it. The thing each of us presents to higher education is a unique case history for the study of change in changing times.

For evaluators to look at us this way will require two things: some understanding of the nature of change itself, institutional and personal, and some appreciation of its relevance to all education, institutional and personal. It is to both of these topics that I would like to speak and to apply my five years' experience in one new and changing college.

I have come to three conclusions about change. None of them is particularly original but they all bear on the problems of evaluating newness and so need to be attended to.

1. I believe that change is as basic for the continuing health of organizations as it is for all living things. Without change, organizations become self-serving, lose transcendent purpose, and benefit neither the persons who serve them nor the society which they serve. The sociology of organizations is clear enough on this. The nature of bureaucracy is to stabilize roles and functions; its ideal is self-perpetuation. It resists self-renewing change. An overbureaucratized world is challenged to discover how to build renewal into organization. Our time of ideological ferment when much of the past has lost legitimacy—including those organizations which represent the past—lacks clear and persuasive alternatives for much that has been condemned. In government, in social institutions such as the family, in religion, and in education, we suffer a continuing crisis of expectations. Learning to live gracefully or at least productively with that crisis—and not regressing to myths of a simpler, better past that never was—is the tough task of all of us, particularly those in education. And the new colleges, where past and future meet with continual urgency, are good places to study this crisis and perhaps learn what to do about it.

2. Not only is change necessary for the continuing health of all organizations, but also for persons. In education, where we deal with persons-in-transition (i.e., students), change is the heart of the matter. If so, then it ought to follow that our educational organizations should be so designed and run as to facilitate human change, growth, and development. But our times have produced a five-foot shelf of indictments of schooling systems hostile to life, to growth, to development, self-actualization, and change. Charles E. Silberman's book, *Crisis in the Classroom* (1970), written with the support of a Carnegie grant, is only the most recent and one of the best of such criticisms of life and death in the American public school.

The new colleges spawned in the '60s were many times inspired with the hope of creating better environments for human development. A study of what they've learned and unlearned would be very relevant to the need of all schooling systems to reexamine their actual impact upon students.

3. The third thing I've learned about change is that it comes in all sizes and shapes. One way to understand the variety is to imagine a continuum. On one extreme you have simple incremental change, what has been called the "Great Society" view of change, which is to do more of what has always been done only to do it better.[1] This approach to change assumes that the essential structure and purposes of things are basically sound but that we haven't begun to do enough with it. Therefore, what is needed is *more* of something. In education it might be more school buildings, or free textbooks, or better-trained (i.e., more discipline-oriented) teachers, and of course always more money.

The opposite view I am going to call "transformative change," although it is more commonly labeled these days as "radical change." If we could use the term "radical" etymologically–meaning "at the root of things"–it would be perfectly adequate for what I am going to describe, but of course our age suffers from a hardening of the ideologies and "radical" is misused left and right. Therefore, in the hope of enlightening discourse, I propose what I believe is the more descriptive, less explosive, adjective "transformative." This view of change asserts that simply doing more within the present schemes of things may not help, may in fact make the situation worse. We may have to escape from our customary way of conceptualizing things in order to re-imagine how to solve a problem. For example, in education, rather than hiring more teachers, we may need to ask some "transforming" questions about the whole learning process: Is the direct intervention of a person called a "teacher" always and everywhere the best or only means to facilitate all kinds of learning? Are "school buildings" always and everywhere the best or only means to stage the process of learning–particularly at a time when many critics see them more often used as warehouses to keep kids out of mischief, off the streets, and off the job markets? Is the traditional ordering of knowledge into so-called "disciplines" and the conventional grouping of faculty into departments functional or dysfunctional for education in a problem-oriented world? The old mathematics of simply incremental change

---

[1] "Great Society" comes from an article by Fred M. Newmann and Donald W. Oliver, "Education and Community," *Harvard Educational Review*, 37, Winter 1967, pp. 68-70. (Much of their reexamination of the premises of contemporary American education implies the need for the kinds of transformative change that I propose in this essay.)

requires mainly the skills of multiplication and division. The new mathematics of transformative change requires that we put two and two together and get curious.

Although change in the "Great Society" mode presents some thorny problems in the old distributive games of economics and politics, it presents no philosophical problems at all. But transformative change, because it pulls at the roots of things, is always an ordeal–psychological, sociological, economic, philosophical, and political. Because many of the new colleges have either asked themselves or have been asked transformative questions, they have often found themselves caught up in all aspects of this ordeal–most ominously, of course, the political. It is one thing to create, say, a new medical college on a campus, thereby changing many things, creating problems of adjustment and accommodation involving faculty, curriculum, and budget–especially budget!–but not really changing any basic conception of means and ends in higher education. Life with father for such a "son of Old U." would present few ideological crises.

But when you create a really new college and freight it with the mission to question the ends and means of living and learning, you deliver it into a crisis that goes beyond the headaches of duplicative curricula, jointly appointed faculty, and competition for scarce dollars. You raise one big and continuing question of legitimacy. Testing everything that exists by asking transforming questions inevitably provokes entrenched interests. It is not as if you were saying, let's improve our football team by better recruitment of athletes, more scholarships, better coaching, finer stadiums, and more effective public relations. It is rather as if you were saying, let's stop playing football altogether because it is basically the wrong game; since its whole competitive ethic requires that one man must be beaten down in order for another to win, it is wrong for the kind of humane society we want to build. That kind of transforming change gets coaches very up tight because it threatens whole establishments.

And yet that is exactly the kind of threatening question that critics have been asking in the '60s. There are no more sacred pigskins. Doing better the same old thing won't do. There are things that are *structurally* wrong with the educational establishment. Those of us inside that establishment can hardly ignore the criticism. Nor can we scapegoat the critics as a red-eyed minority. We do have the choice–and the necessity–to see what is valid and what is invalid in the criticism being made, and the opportunity–particularly in the new experimental colleges–to set our houses in a new order. Most likely it will be there that we can learn best about the consequences of transformative change.

Change as transformation seems to involve at least two typical crisis stages. The first centers around the crisis of unlearning the past, the second around the crisis of inventing alternative futures. Both are crises because both put some portion of the presumably stable, known, and secure world up for questioning. Yet it is a commonplace insight into the creative process that both crises are necessary if we truly entertain the desire for change, whether we are talking about a room full of brainstormers or a college full of rain makers. The destruction of a previously accepted scheme of things must precede the invention of a significantly new one. Or in more drastically metaphoric terms, transformative change requires that we die to an old life before we can be born to a new one. Unless we are willing and able to do that kind of dying, we are often condemned to experience another kind of death which students of this generation have come to call "irrelevance." Whitehead didn't use their term but he referred to the same problem in education when he said that the secondhandedness of the learned world was the secret of its mediocrity. Learned conventions, as William J. J. Gordon (1961) writes, can be windowless fortresses which exclude viewing the world in new ways (p. 92).

The real story of many of the new colleges might be summed up so far in what they have done with this first crisis of unlearning. Inevitably it seems, whether we are seeking to transform an organization or ourself, the first things we learn are usually negative, rejections of some tribal pieties of the past. We discover who we don't want to be and what we don't want to do before we find out who we want to be and what we want to do. It's true of most undergraduates. It's also pretty much the story of Justin Morrill College (JMC). We have been divesting ourselves of a good deal of the conceptual baggage packed into our guidelines by our solicitous parents when we were set off on our journey five years ago. But we quickly discovered that this is an age that is more often than not beyond its maps. Our inherited charts were often unrealistic or irrelevant to the actual undiscovered country we found ourselves in. We have been trying ever since to get out of the world of discourse that ties education up into categories like "teacher," "classrooms," "courses," "credits," "degrees," "grading," and "disciplines." Our success so far has been limited.

For one thing, it takes guts to sail off your maps. For another, it takes time. If one problem of the first crisis of transformative change is legitimacy, one problem of the second crisis—to invent alternative futures—is time. Time to get a possible idea. Time to shepherd it through the approval system of a status quo. Worse than merely "taking time" though is your own apparent inefficient use of time. Your days are full of false starts, crazy schemes that don't work, seductive ideas that fall apart, many trials and many errors. Vermonters who invented maple syrup learned something of the mathematics of creation when they found it took 60 gallons of sap for one gallon of the pure sweetness. Unfortunately using time just to play around with ideas flies in the face of the American ethic of practicality. It requires something of the Eastern wisdom: "Don't just do something. Stand there!" But critics, friendly and otherwise, are more often than not bewildered by what appears to be a pattern of inconstancy, instability, vacillation, irresolution, whimsicality, and faddishness. As Emerson warned us long ago, for most people, consistency is a hobgoblin.

As I read summaries of the results of new colleges so far, for example Jerry Gaff's book on *The Cluster College* (1970), I am more impressed, as I believe he is, by the questions they have raised–i.e., the unlearning they have done–than by the answers they have given. And that goes for Justin Morrill, too. It is much too early in the game to look for results in terms of either tested or even testable models, but it would be very much to the point to evaluate what such colleges have unlearned and why.

I know of course that the universities which created us won't support for long the only kind of answer most of us can honestly give at this point: "We've been asking some mighty good questions!" Although we may think that that's "doing something," those who supply dollars and legitimacy may believe it's only "standing there." For similar reasons, evaluators are often unhappy with us, too. In their eyes we fail to "define our objectives," we seem perversely to insist upon using vague and "untestable" language about human development and self-actualization, we no more than get started down one road when we take off on another, leaving the tracking evaluator bewildered as to what we're doing or even where we're going. The evaluator begins to suspect that the innovator is being deliberately secretive because his ego can't stand the thought of failure. On his side, the innovator suspects the evaluator still believes the old maps adequately describe the new world. He can't see the questions for the answers all around him. But we are often convinced that the traditional measures used for evaluation are such crude nets that many if not most of the important fish get away. Often we also suspect that the evaluator is so defensive about his tests that he won't admit the reality of anything they can't measure: "If I can't whistle it, it ain't music!" Worse, as innovators we often look at the evaluator as the servant of the status quo, one who measures all differences against their sameness.

I'm not so politically naive that I don't fear the politics of change. Anybody in the change game will need all the cunning he can command just to stay alive. As a politician for change, I believe that existence precedes essence—and this has absolutely nothing to do with Sartrean existentialism. It simply means that with existence, you can, perhaps, do something. Without it, nothing. So when in doubt, I punt. When necessary, I compromise. When required, I conform. But beyond the politics of survival, which require that distrust, competition, conflict, and suspicion be blunted, appeased, papered-over, and placated, there is another arena where just the opposite must be true. In the philosophical debate over transformative change, we should let it all hang out. There the conflict between evaluator and innovator should be open and honest. Each needs the other's frankness to preserve him from the defects of his own virtues. The strong suit of the innovator is his willingness and sometimes his capacity to see things with new eyes, to convert answers into questions, to call what is to account, to discover that the emperor is naked. His defects are almost always vagueness, constant shifting, contempt for management, and a congenital skepticism for dry-eyed analysis. The reverse of these is, of course, precisely the virtues of the evaluator who insists upon defining a clear target in order to measure, quantify, manage, and regulate.

I'll accept the evaluator's reasonable and insistent—if often irritating—challenge to define goals and objectives if he will accept my premise that in our times we are beyond the maps and so we will probably have to come up with new charts. The challenge of being evaluated is a good one. Every organization should face it regularly, and with its very existence on the line, as is often the case with the new colleges. As I have argued earlier, organizations must learn the fine art of dying. I hope I am being honest, if a bit dramatic, when I say that we in JMC are prepared to die—as an organization, you understand—and to die willingly for the sake of a better idea. We've been practicing dying to old ideas the last five years. After all, the concept of a schooling system we all share—the "college"—and in particular the concept of a "residential college" is an ancient piece of tribal apparatus that may be no more relevant for a university in the year 2000 than the saber-tooth curriculum or the department. My only request in this dying business is that we die usefully. Merely being eliminated because we are a threat to someone else's living isn't good enough. Being found irrelevant, outdated, or unproductive is.

To the evaluator I would say this. I don't know what's in your literature on the subject of change. Everything I've learned may all be there. But as in so much research on education, it may also be only there, packed away in the library, and as yet unused on the playing fields. In my experience in the change game the last five years, I have not seen enough evidence of the kind of change-conscious systems-wide evaluation of new (or old) colleges that I would desire. My inference is that much evaluation stays within the inherited universes of university discourse—"classes," "courses," "grade point averages" and so forth. In Justin Morrill we have taken many "snapshots" these past five years. We have administered the Graduate Record Examinations to seniors, we have given seniors a structured interview covering many aspects of their four years, we have compared grades given in the college to those given in the university, we have tested our language students against Modern Language Association proficiency measures, we have administered CUES, and CSQ, and CQT and OPI, we have given freshmen the Myers-Briggs indicator, studied the effects of our overseas experience on "worldmindedness" and so on and on.

Although all of these give valuable information about our various programs, no one of them and no combination of them steps outside the charmed circle and asks transformative questions about our experience with change. Would I be too fanciful if I proposed some new measures? For example, how about a "risk scale" to measure not how much has been chewed but how much has been bitten off? A "shoreline scale" to see how far out we have waded? A "question scale" to determine the number of things we have converted from answers to questions? A "wastebasket scale" to see what we have found won't work and why. An "awareness scale" to see how broad our horizons have become as a result of being immersed with change. And maybe even an "invention scale" to total up the sheer number of new ideas we have teased out, whether good, bad, or indifferent. Above all, I would like to see a systems-wide systemic approach to get at the interactions of a total, evolving organization.

There are, at any rate, numerous straws in the wind that the evaluating/testing game is changing. For example, according to a *New York Times* story on November 1, 1970, a special study group appointed by the College Entrance Examination Board has criticized the board's national testing program for focusing mostly on the interests of colleges but failing to serve the interests of many students. The group asserts that board examinations such as the SAT fail to recognize and assess a wide variety of talents, skills, and mental attributes, thereby penalizing many different kinds of students—or rather, many kinds of "different" students. As a result of this three-year study, the commission urged extensive reform of the testing program. In particular, I was struck by the group's recommendation to develop a detailed "test" that colleges themselves, as institutions, would take. It would provide students with a more accurate basis for choosing a college. The idea intrigues me, but the suggested examples as to what such a test would embrace disturb me: statistics on class size, time spent in class by professors, the faculty-student ratio, the number of students who drop out each year and why, and a measure–obtained from a student questionnaire–of the school's social and intellectual climate. Except possibly for that last idea, all of the preceding suggestions simply accept the present scheme of things within which to measure the college, for example such things as "classes," "professors," "student-teacher ratios," and "dropouts." For the latter the simple substitution of a new name such as "mind changer" to replace "dropout" might help to re-cast the entire situation in a more productive, less stigmatic, light. So long as we continue to measure colleges this way, or by the numbers of graduate degrees we produce, we will never see whether the emperor has any clothes on, but only the fashionable cut of his fancied suit.

In Justin Morrill we would like to have a college test run on us that would focus on two areas in which we think we have made a beginning: (1) facilitating transformative change of our institutional setting itself, and (2) facilitating a set of attitudinal competencies related to liberal learning of our students for changing times.

Because we believe that change is necessary for the continuing health (i.e., relevance) of the organization itself, we have courted change of Justin Morrill as a schooling organization since we began. It is not appropriate here to detail the number and kinds of changes we have made in JMC the last five years. Changes began, however, from the moment the first faculty committee was assembled to plan an actual program within the guidelines delivered to us from the university. These guidelines were open enough, and–because of the inevitable cross-currents of ideologies embraced within them–confusing and contradictory enough not only to permit but absolutely require interpretation, modification, and transformation. Our mission, we thought, was not to create an unchanging model from a guideline, but rather to provide a staging area where old ideas could be dismantled and new ones put together. To me then as now it was important for JMC to become what I call a "trying college" in all senses of that word. Our changes have ranged from big things to little things. Most are incremental, doing more of the same old thing (for example, different numbers of credits and hours for courses). A few approach transformative change. One example of the latter is our so-called "wild card" curriculum. Rather than requiring a common experience of all students (and all faculty) in the usual general education year-long courses, we instead let teachers teach anything they wish from the point of their greatest interest and we permit students to choose from this great variety of individual courses those which interest them for whatever reasons. Although we do set certain demands of distribution to assure breadth, we have introduced throughout the student's four years a great many points where he must make his own decisions. And we have put a good deal more interest into our classrooms on both sides of the desk. Obviously such a system assumes different outcomes more related to general competencies than to a common general education knowledge.

Our hope is to encourage the development of a certain set of intellectual and attitudinal competencies. With our community involvement system called the Forum, our "learning coach" strategy for teaching, our administrative style which tries to say "yes" to the possible, we hope to facilitate such competencies as open-mindedness, the capacity to ask transformative questions (i.e., to unlearn), a tolerance for ambiguity, modes of awareness beyond the cognitive, problem-solving, and system-building. Such competencies, although relatively uncommon in lists of traditional academic learning, are nevertheless more and more frequently being cited by those who study the future as being relevant skills for coping with change. Although each of these competencies may lack the necessary definitional "hardness" to fit easily into an evaluator's measure, each may nevertheless present a relevant challenge to the evaluator to develop measures appropriate to the times and to the human attributes necessary to cope with them.

I am not of course saying that Justin Morrill–or any other college in the change game–has as yet developed much that is tested or even testable. Our actual colleges are bundles of compromises between past and future, prettier in prospect than any piece of the action will often reveal.

To conclude: there are different ways of keeping score in the change game. The innovator and the evaluator must understand that they play on different sides, but that both are necessary to win anything worthwhile. Sidney Cox in a book about the creative process called *Indirections for Those Who Want to Write* (1962), put it this way:

> We all live in two realities: one of seeming fixity, with institutions, dogmas, rules of punctuation, and routines, the calendared and clockwise world of all but futile round and round; and one of whirling and flying electrons, dreams, and possibility, behind the clock [p. 127].

The innovator lives more often with flying electrons, the evaluator with rules of punctuation. It is the combined task of both to help transform some of the primordial flux into the provisional fiction-fabric we call the clockwise world.

References

Cox, S. *Indirections for those who want to write.* New York: Viking Compass Books, 1962.

Gaff, J. G., & Associates. *The cluster college.* San Francisco: Jossey-Bass, 1970.

Gordon, W. J. J. *Synetics.* New York: Harper and Brothers, 1961.

Silberman, C. E. *Crisis in the classroom.* New York: Random House, 1970.

Chapter 7

TOWARD A STRATEGY OF EVALUATION FOR A NEW COLLEGE

Beth B. Mason
Douglas R. Moore

In September 1967, Callison College of the University of the Pacific began operation with 80 freshman students, eight faculty members, and two administrative officers. The first provost had been provided one year lead time to recruit the faculty and staff and to organize the curriculum. The maximum enrollment was projected at 300, with about 85 of these students studying abroad each year. Ultimately, the plan called for 21 faculty members and two administrators on the Stockton campus.

As with the two other cluster colleges at the University, Callison's program was developed around a distinctive motif, that of nonwestern studies. The fundamental idea has been that experience with and in a nonwestern culture would make a significant impact upon the student by providing him with both motivation for a more thorough and critical understanding of his own roots and a wider world view which would be a foundation for international understanding.

Each Callison student will be required to spend his sophomore year living and studying at a center in Bangalore or an approved alternative. The expectation was that should the general education program and the year in India have a significant impact on the student, he should have the final two years in which to adjust his studies to his new insights. This plan is still preferred, but experience has shown that, for various reasons, a few students postpone the year in India to their third year.

Although the philosophy and goals at Callison were generally more tacitly accepted than expressly articulated, there were some basic assumptions that received wide and continuing discussion. Among these were four of especial significance:

1. *Change.* Nothing was considered sacred. Openness and a flexible structure which would enable changes to occur easily and even to a radical degree were considered as a necessary condition for the development of the college.

2. *Self-criticism.* From the beginning, "navel gazing" was virtually a way of life at Callison. The faculty and students alike talked often of the need for self-criticism and the ongoing process of self-evaluation.

3. *Pragmatism.* Relevance in education means that there are social correlates for the intellectual conceptualizations. If what one studies is relevant, it has application to the larger society. A part of this practical impetus was the commitment at inception to a program of evaluation.

4. *Commitment to developmental goals.* Callison self-consciously stood under the influence of educational researchers such as Katz (1968) and Sanford (1962, 1967). There was consensus on the importance of both the cognitive intellectual dimension and the affective developmental processes of college youth. Further, evaluation of such programs derives from what qualitative changes occur in the student between the time he is admitted and the time he leaves the program. Callison, from the beginning, rejected an elitist philosophy whereby the evaluation would derive a positive connotation from a high profile on CEEB scores or the high school grade point average (GPA) of applicants.

Role of the Original Faculty and Charter Class

The faculty ranged in age from 27 to 41 years. All had previous teaching experience; four came from within other units of the University. Research was of minor significance and all the faculty wished to place their major emphasis on classroom teaching. Only two would be considered "India specialists."

Many foreign area study programs emanate on campuses where the missionary impetus has been, or is, strong (Mayhew, 1968, p. 7). Among Callison's initial faculty 50 percent held theological degrees, 20 percent were children of missionary parents or had previously held strong commitment to institutional religion. None of the faculty was a practicing churchman and only one faculty member retained any identification with his theological vocation. The influence of these ex-ministers upon the development of Callison's program was described by one faculty member as an infusion of "post-Christian missionary zeal."

Many of the faculty had read Nevitt Sanford's recent publications on collegiate issues and were in agreement with his views. There was consensus with the statement that, "Instead of judging a college's success by who goes there and how well they are expected to do, I want to suggest that we focus on what happens at the college, on what kinds of change it induces [Sanford, 1967, p. 19]." Although most of the faculty were products of large universities and traditional graduate schools, they agreed to the suitability of the challenge that "The measurement of change offers a fair test of what a small college or a new college can do [Sanford, 1967, p. 20]."

Finally, a defining quality of the original faculty was its anticompetitive bias. The college operated on an egalitarian policy in which everyone, including students, was involved in making policy. The administrators were also teaching faculty members. Competitive attitudes were further reduced by agreement upon a pass-fail system of grading and a policy of participating in no university programs which would make invidious distinctions between students (such as honor societies).

To the task of innovation, the faculty at Callison brought a greater-than-average concern for value orientations and a skepticism regarding institutionalized solutions, a commitment to effecting change and working in an egalitarian relationship with contemporary students. The researcher's problem became that of discerning these values and orientations in the face of the faculty's reluctance to articulate personal views for fear of interfering with group process. In their insistence upon consensus, they were acknowledging that "the history of higher education shows that new innovative colleges are seldom more experimental than in their first period of idealism and energy, after which changes tend back toward tradition. This is another reason for cluster colleges to reach for an identity based more on dissimilarities than on duplication [Martin, 1968, p. 76]."

The charter class of students was also crucial in Callison's subsequent development. The decision to place themselves in the role of "guinea pigs" in a new college defined the students as innovators. They were seeking experiences not yet defined in their minds, willing to tolerate the uncertainties and risks involved with innovation, and seeking to share in shaping and defining their educational context.

The qualities of the charter students which were relevant to the role in which they placed themselves may be subsumed under three headings. First, rejections were more clearly defined than commitments. They rejected expertise. Although the majority maintained an A or B average in high school and a high percentage had been members of scholastic honor societies, few expected to graduate from college with honors or with a high GPA. Although their parents were a highly literate group, far above the national average in postgraduate degrees for both parents, the modal category for anticipated degrees was only the master's degree, and few students anticipated doctorates of any kind. They rejected positions within the Establishment. The charter class already had considerable organizational experience and political acumen, but the percentage anticipating positions of community leadership was small. They rejected conventional ideational stances. Few considered it important to be well off financially or be a success in their own business; a high percentage reported themselves as "religious dropouts" (religious background data compared with current religious preference); and, finally, many rejected socially defined stereotypes of the masculine and feminine role.

Second, they were simultaneously altruistic but profoundly skeptical. A high percentage ascribed to goals of social amelioration (help others in difficulty, join Peace Corps/Vista, preferential treatment for disadvantaged). On the other hand, the students were skeptical of conventional channels for achieving these goals. They were committed to change but the direction of that change was not clear to them. Much of the ambivalence which they experienced was related to developmental conflicts which were apparent in personality test scores. But the personal conflict was projected into the students' doubts as to the relevance and ultimate meaning of commitment to any goals. The striving for power and ascendancy shows clearly in low dependency needs, high needs for Self-Assertion and Emotional Expression, and (for males only) a low sense of Well-Being. Communication interests and skills appear to be the students' common point of definition rather than the social sciences. Vocational interest scores were highest in Fine Arts and Communication occupations for both men and women.

Finally, if there was a commitment among students, it was a commitment to change and novelty. Personality test scores showed both males and females to be high in Flexibility. In spite of acceptances to traditional collegiate programs, they elected a novel, largely unknown situation. Yet they came to college prepared to participate in demonstrations, ready to engage in a power struggle vis-a-vis the faculty, and in general both to assert themselves and to place themselves in a position to be challenged.[1]

Curriculum

Innovation and experimentation in higher education is largely a relative matter. The Callison staff was aware that they were attempting little which was actually new. The novelty of the Callison program is attributable to three elements: the setting in a quiet, traditional university; the requirement of a year spent studying in a nonwestern culture; and the particular constellation of program elements which were integrated (the curriculum and faculty, the curricular extensions in the form of projects, and the psychology of student life).

The original freshman curriculum was comprised of four courses running both semesters and required of all entering students. The traditional semester calendar was modified to incorporate a month of independent study during January. (The so-called 4-1-4 Calendar.) The curriculum was:

1. *Heritage of Man.* This team-taught course was designed as an examination of the traditional values held by western man. These values were juxtaposed with the contemporary social and intellectual issues and crises in the experience of western man.

2. *Socioeconomics.* This course was concerned with both historical and current social and economic problems. Gradually it came to focus on issues in developing societies, using India as a case study.

3. *Scientific Thought.* This course was concerned with the history and philosophy of science and an exposition of scientific methodology.

4. *Anthropology.* This course dealt with history and origins of culture as well as methods of comparative cultural studies.

5. *January Independent Study Period.* This study was originally designed to accomplish two things: initiate the student in independent methods of inquiry and to give orientation for India since his topic had to deal with some facet of Indian life or history.

---

[1] Beth B. Mason, unpublished data, 1968-70. Permission to cite granted.

A unique requirement for the freshman year, carrying no academic credit, is the Community Involvement Project. Each first-year student must spend a minimum of three hours per week working in one of the subcultures of Stockton. Students have been involved in such settings as the California Youth Authority, the State Mental Hospital, Head Start Programs in the public schools, and Homes for the Aged. The rationale for such a requirement was two-fold: to sharpen the consciousness of social issues and to provide an opportunity to develop facility in dealing with persons with a different life-style. This program has been strengthened by employing a full-time director. Many third- and fourth-year students voluntarily engage in the program.

The original faculty was recruited to develop and man these courses. Since the first two years of operation would each have only one class of students on campus (the second-year students being in India), a two-year trial period for this curriculum was assured, but student evaluation of the curriculum and student participation in the decision-making process created change.

Freshman requirements were reduced to the *Heritage of Man* and *Socioeconomics* courses. *Socioeconomics* was changed to a team-taught format and broadened to include more of a social science methodology introduction. *Scientific Thought* and *Anthropology* would no longer be required, although a minimum one-semester course in a natural science must be completed at some time before graduation. The January independent study remained a requirement but need not focus on India.

Pursuant to the conviction that learning brings its own intrinsic rewards, no grades were issued to students, although grades were actually recorded. The student was only to learn of the grades he had received when he graduated. He received detailed letters of evaluation at the end of each semester and notice of pass or fail. After two years of this system, both faculty and students felt this to be an insidious method creating unnecessary anxiety. The college then adopted the simple method of recording only courses satisfactorily completed. This record contained the single designation of "credit." The letters of evaluation were continued.

The program in India, manned by Indian faculty, included in the beginning Indian Art History, Politics, Religion, Economics, Literature, History, Language, and the Arts. The changes occurring in that program were in making the language optional, increasing the offerings in the Arts, and providing more extended periods for travel. All these changes were initiated by evaluations made by students. This program was operated for the first three years from the facilities of a small hotel near Bangalore University. During the fourth year, the study center and the living-dining facilities are being moved to other quarters better suited for educational purposes.

With the beginning of the third year, a new set of dynamics was introduced. The first class of students returned from India and for the first time there were two classes of students on campus. This was also the time when the first significant changes were made in faculty and administration. Seven new faculty members were added and the original provost resigned to start a new college in Indiana. The first preceptor (second administrative officer) was elected as provost. The addition of new faculty permitted the development of the curriculum for the junior and senior years of study as well as strengthening the program for freshmen. Emphasis in hiring of new personnel continued to be placed upon (1) competence in an appropriate field of study, (2) experience and interest in nonwestern study, and (3) a commitment to interdisciplinary education. The nonwestern dimension was restricted to South Asia, East Asia, and Southeast Asia. Studies in these areas would be on a comparative basis, particularly in reference to western culture. Students would major in either social science or humanities with a minor in international studies.

A new international element was added to the program. Students demonstrating interest in anthropology were provided opportunity for a semester of field study in the Yucatan of Mexico. The first group of sixteen students was in residence in Yucatan in the spring semester of 1969-70. This first group tended to support the hypothesis that a year of study in India will provide the student with both personal and intellectual skills relevant to successful functioning in another alien culture.

Since the structure of the program in India demands a great deal of autonomy on the part of students, the first year at Callison provides opportunity for the student to experiment with and develop his skills in assuming adult prerogatives. Inability to handle freedom in an adult or mature manner could prove embarrassing for the program in India and/or damaging to the student himself. These concepts and the evolving pragmatic tests determined both the original curriculum and its subsequent revisions. Obviously these concepts were determinative in the recruitment of faculty as well.

Varieties of Evaluation in a New College

There is a certain incongruity when a college dedicated to innovations attempts to evaluate its efforts against standards established for traditional institutions. Indeed, as Miles (1964) has observed, "educational innovations are almost never evaluated on any systematic basis" but by substitute methods such as sentiment, educational ideology, persuasive rhetoric, or intuition (p. 657).

Belief in the uniqueness of a particular educational experiment, an almost mystic reverence for the combination of factors which joined to conceive the experiment, often results in the belief that an experiment cannot be evaluated. Like the proud parents of a first child, innovating faculty often feel that no outsider can properly appreciate their progeny. Nevertheless, the reality test is inevitable.

> The issue is not whether an educational program is to be evaluated; rather it is simply a matter of the kind of evidence used in doing it. . . . in the long run, judgments made on the basis of relevant evidence are more valid than subjective emotional reactions. Yet, beyond the objective evidence no matter how extensive it may be, are values and practical considerations which must color judgments [Dressel, 1958, p. 19].

Prompted by a self-conscious interest in its own development and the social science bent of its faculty, the Callison community proceeded from its inception with an almost continuous self-evaluation. The University provided formal data from admissions records, such as SAT scores, high school records, etc. The Counseling Center administered a variety of test instruments at regularly specified intervals. Students used questionnaires and interview techniques to study special aspects of their collegiate program and themselves as an experimental group. Faculty members collected responses to the India experience and student evaluations of each class offered. But the most extensive source of data was the intimacy of the college itself. Student development, results of curricular experiments, success of dormitory governance—all were quickly visible to the entire community and were the subject of intensive discussion. Perhaps no experience other than that of living through the initial years of an experimental college could provide the researcher more convincing evidence for the contention that "The most persuasive evidence comes through a triangulation of measurement processes" (Webb et al., 1966, p. 3).

Research Questions

The research projects described above might be subsumed into the following research questions:

1. *What kinds of students elect a college such as Callison?* Is it the curriculum, the experimental nature of the program, the year in India, or

some other unknown combination of factors which attracts these students? Given the probability that the students are a distinctive group, what are their special needs? What are their goals? Finally, who will prosper in a community such as Callison provides? Who will be disabled by the experience? Who will drop out? Experience and empirical test data have demonstrated that these are distinct groups with little or no overlap. "Dropping out" has not proved to be a measure whereby it was possible to distinguish the college's successes from its failures; it proved to be an overly simple measure.

2. ***What are the developmental stages of students in this program?*** In a traditional program, it is relatively simple to distinguish between lower-division students and upperclassmen. Given a stabilized collegiate program, an experienced faculty member can usually distinguish a freshman from a sophomore, or a junior from a senior simply on the basis of his work, demeanor, and central concerns. Not so in the innovating college. With the exception of qualifying for and surviving the India year, there are few hierarchal guidelines apparent. Both faculty and students need a conceptual framework against which to compare a student's development and measure his progress. This is particularly true in the absence of a traditional grading system.

3. ***What should the India year provide?*** The Callison program is deliberately oriented to general education goals, not to the development of India specialists. It is a fine line to draw, however, between those experiences which provide the student with adequate preparation for his encounter with an alien culture–experiences which allow him to be a knowledgeable student rather than simply a tourist–and those which are more appropriate preparations for a career in India specialization.

Many of these questions cannot be answered definitively before the college has even graduated its first class. But innovators–be they students or faculty–are not noted for their patience. So data continue to accumulate and research projects develop providing the material out of which more definitive hypotheses and conclusions will be built and stated.

Data Sources and Implications

In many ways, Callison students do not differ from other University of the Pacific students. Admission to the University suggests that the students had achieved a certain level of academic proficiency, came from a relatively affluent home, and preferred a private school to a state college or university. College Board scores indicated a mean of 563 Verbal, 523 Math, and standard deviations of 138 for composite mean of 1086. Females scored slightly higher than males. Sixty-one percent of the students had higher scores on the Verbal than on the Math section. These data from the University admissions office provided few clues to the distinctive qualities of the Callison student.

Data from the American Council on Education (ACE) survey provided more extensive information on the entering freshmen. The profile emerging from these data has already been presented. Clearly these students were marching to the beat of a different drummer. The themes which emerged from analysis of ACE data were those of social concern and involvement along with openness to change.

Panel design and test instruments. The second major source of evaluation data came from test data collected according to a panel design. Incoming freshmen were administered the College Characteristics Index–CCI (using "expect" instructions), its companion the Activities Index (AI), California Personality Inventory (CPI), Strong Vocational Interest Blank (SVIB), and a questionnaire of demographic data. In later classes the Omnibus Personality Inventory (OPI) was substituted for the CPI and AI. Each student was later given a packet containing his test scores and a profile of the norms for his class and an opportunity to ask questions of the examiner, seek further clarification, etc. At an explanatory session the faculty was presented with group profiles for each test.

Immediately before departure for India, these same students were retested using CCI and AI.

Following their return from India, the students repeated one of the test instruments, usually one relevant to psychological development–either the CPI or the OPI.

Toward the latter part of their senior year, the members of the charter class will retake most of these tests–the plans being to include an instrument descriptive of the college, an instrument descriptive of the student's self-concept or psychological dimensions, a vocational interest test, and a questionnaire of items most relevant to change of attitudes, values, and objectives.

The panel testing provided results which were at times immediately relevant, at other times led to theory building, and at times simply provided amusement. Group norms on the SVIB clearly indicated that the students had an interest in the Fine Arts equal to their interest in Social Service–a dimension that had not been anticipated in the curriculum. Of use in long-range planning and theory building was the unexpected finding of high flexibility scores on the CPI. The faculty had expressed much concern over whether the students would be able to make the transitions necessary for participation in the India program, whether they would be able to adjust quickly to an alien culture without serious psychological disturbance. As it turned out, males achieved a mean standard score of 58.03 and females scored 65.03. These scores along with observational data led the faculty to conclude that for many students, not rigidity but excessive flexibility was a problem and the educational need was to provide more structure, if not rigidity. There were also amusing and unanticipated results from the testing program. The tests had been selected in part on the basis that their results would be interpretable and useful to the students. On the basis of the social service and fine arts bent of the students, means on the Masculinity-Femininity scale (SVIB) might have been anticipated. But it was an enlightening experience for the researchers to explain the low Masculinity scores to a group of irate, defensive freshman males who were not about to have their virility impugned.

Other sources of empirical data. One of the richest sources of data, and most immediately useful, proved to be the student evaluation of faculty and/or classes. The agreement to reciprocal evaluation seemed to remove some of the acrimony that often characterizes student-directed evaluation programs.

Little research has been possible on the India year, although it is a subject of major concern to all members of the Callison community. One visiting faculty member, however, using a simple questionnaire technique, uncovered a remarkably predictive yet heretofore unnoticed fact. Students who do not like Indian food (and they have extensive exposure to it before electing to spend the year in India) are the most likely to fail to complete the year abroad.[2]

Inasmuch as the faculty and student bodies are still small, interviews provide one of the simplest and most direct sources of information. For example, within the first term of the college's operation, each faculty member was asked to cite the names of a few students who were prospering in the innovative conditions and a few who should not have come to an experimental college or who were ill-suited to the program. Each faculty member was asked to explain the basis of his nominations, and from these observations it was possible to extrapolate the following data. Students most likely to succeed in an innovative situation have the following qualities:

---

[2] Patricia Schedler, unpublished data on Callison students in the India year, 1969-70. Permission to cite granted.

1. *Endurance, both psychic and physical.* Successful students were described as unwilling to give up easily, as having sufficient energy to go outside the college community and get involved in non-Callison activities, as demonstrating sufficient stamina to drive toward a level of excellence, regardless of initial difficulty, and as willing to encounter a new situation.

2. *Good coping ability.* They structure their own world by managing time and energy well and are not threatened by assignments whose parameters are not specifically defined; they seek out new experiences and are actively coping with *present* challenges rather than being hung up in psychic cobwebs of the past.

3. *Able to participate in and create fun.* They discover the joy of inquiry and excitement over new ideas, see humor in stressful situations, and relax in laughter, do not take themselves too seriously, and are "happy, excited and exciting people."

4. *Possess qualities of openness.* They are willing to be challenged and respond with openness and intellectual flexibility, are able to question personal values without being unduly threatened, and are generous with themselves and their resources.

5. *Willing to learn.* They are willing to attempt new styles of communication and substitute them for styles in which they have already been successful, are willing to work through problems or issues even if they cannot be resolved or lack an absolute "answer," are either confident in the powers of their own intellect or just beginning to discover their own powers, are able to be creative or inventive in the handling of original sources, and some–though not all–are already well trained in critical thinking and widely read.

Students viewed as being negatively affected by or unsuited to the program were described in the following categories:

1. *Devotees of "instant wisdom."* These students were described as "authoritarian lefts" or "hippies who want to relate everything to mystical knowledge without the aid of traditional assistance, like books" or "students who are looking for a philosophy of life without having to deal with nitty-gritty issues."

2. *Neurotics too well defended.* They were described as being aloof, refusing to become involved or experience either excitement or threat. Some were comparatively unstructured personalities whose attitude conveyed "everything is relative; nothing is worth getting excited about." Others were narcissistic, conveying "don't teach me anything that might upset my creativity."

3. *Cases of "poor fit."* Into this miscellaneous category fell students who could not provide their own discipline and needed a more structured situation, students who were desperate for acceptance and too poorly defended to achieve any degree of autonomy, those students who were prisoners of the past and had little psychic energy to invest in current activities, and finally, those who could have prospered equally well and with less effort in a more conventional setting.[3]

This simple research is cited merely as an example of useful information which may be simply gathered. Despite the inevitable validity problems inherent in such data, four years of experience have provided ample substantiation of these early observations.

Visibility as a source of information. It is the major contention of this paper, however, that so long as a college remains of a size where students and faculty can know each other by name and are committed to mutually agreed upon goals, the best source of evaluative information is provided by visibility. Results of one's efforts are quickly visible and are the subject of discussion among a sympathetic community. The de-emphasis upon competitive career striving among the faculty probably accounts for the lack of defensiveness which allows for such evaluation to proceed spontaneously. As the student body and faculty become larger, it is inevitable that the degree of intimacy and unanimity of the early years will diminish and that the uses of visibility will lessen and need to be replaced with more formal evaluation procedures.

The practice which best illustrates how evaluation proceeds through informal channels in a small innovative setting is that of the screening procedure for the India year. Since the charter class first began its preparation for the year in India, the faculty recognized the need for a process of screening students before they were approved for participation in the Bangalore Program. A full year in a foreign culture 12,000 miles from home may be too demanding on some sophomores. Three criteria were used in the screening process:

---

[3] Mason, op cit.

1. Academic record–a student must have demonstrated motivation for and ability in meeting satisfactory academic performance.

2. Physical health–no student should be approved for the year in India who does not evidence the physical strength and stamina to live in India without danger to his health.

3. Personal maturity–a student must demonstrate qualities of personality development which indicate that he will be sensitive to customs of the host culture and not be an embarrassment either to the Indian people or to the College program.

The third of these criteria is the most crucial and the most difficult. The facilitating factor has been the small and intimate character of the college which provides a high visibility factor. Multiple opportunities continuously are provided for observing students in a variety of contexts. Three of these contexts have been considered more significant than others.

1. The close social situation of the students, the small number, the coeducational residence halls, and the many personal contacts between members of the community give ample opportunity to observe the behavior of students, their autonomy, ability to establish personal standards and to act in a manner sensitive to others.

2. The Stockton project provides opportunity for the student to give evidence of his ability to initiate and follow through on a project commitment involving persons different in life-style and background.

3. The January term of independent study demands the capacity for working independently, without sustained supervision.

These contexts and the qualities necessary for successful participation are assumed to be significant indicators of success potential in India. The high visibility factor at Callison makes possible their utilization. The final decision on participation in the India program rests with a committee comprised of four faculty. In the fourth year, two students who had been to India were added. Until a greater reservoir of data has been collected and analyzed, the information gathered from test instruments cannot be utilized in the screening process, but as correlations are established and hypotheses developed, these data may become integral to the work of the screening committee.

Callison College, Second Phase

Clearly the college has moved into the second phase of its evolution. In June of 1971, the first full cycle will be completed with the graduation of the charter class. Evaluation of the college has to this point been almost totally formative; i.e., rather exclusively in reference to development of the program. It is time to begin the integration of these observations and to draw inferences from the first phase of the life of Callison College.

Problems

That there are difficulties and potential dangers in the process of conceiving and developing a cluster college seems almost too obvious to demand articulation. Indeed, there are times when those involved in such an enterprise thoughtfully observe, "It would be nice to relax in the security of a more traditional situation." Whether there is security in a more traditional situation may be a moot question, but the day-to-day tension resulting from forming a new and innovative institution often makes tradition appear as an ideal state of affairs. Working without benefit of precedent or codified regulations, dealing with virtually every issue on an ad hoc basis is a demanding situation. The expenditure of time and energy, both psychic and physical, is a factor all too apparent to everyone who has gone through it. This is precisely the point at which formalized evaluation may be of value in the analysis of past decisions and in establishing precedents which will be useful in future problem solving.

There are both external and internal problems in such an endeavor as Callison College is pursuing. While this chapter is dealing largely with intracollege evaluation, mention must be made of two external dynamics which are potential problems.

The first of these is relations with the parent or sponsoring body. Questions of autonomy in program development, of equitable allocation of budget resources and facilities are but a few of the many potential trouble areas.

The second external factor which easily can dissipate energies and time is the interest directed toward the project by individuals and institutions quite unrelated to either the college or its parent university. Scarcely a week has passed since 1967 that some individual or team of inquirers has not visited

Callison College. In addition, inquiries, questionnaires, and even test instruments arrive by mail every week. Callison has been included in many surveys studying new, innovative colleges in general and cluster colleges in particular. The aftereffect of this experience has been unfortunate. In only two cases have results been returned to the college. One of these was a superficial summary of the "joys and excitement" of such an adventure. The other was a computer printout without interpretation or even explanation of its format. Such experiences are not conducive to participation by faculty, staff, or students in the evaluation processes. As one student phrased it, "It's like living in a zoo."

It is, however, within the institution that the critical problems and conversely the imaginative opportunities are present. The most pervasive factor holding the potential for both problem and opportunity is that of an emphasis upon affective experience. When a college seeks to balance the cognitive and the affective, to give attention to the developmental needs and tasks of its students, it faces a number of potential dangers. Emotional needs are thrust into the total educational arena. An issue is never purely intellectual; it always holds the opportunity for expression of and attention to "feelings." The following are potential dangers inherent in such an environment.

1. Given the contemporary mood of college youth, an antiintellectual bias will continue to be a threat. "Experience" may become the "thing," with an undue emphasis upon group process. Most students have entered Callison from a secondary school experience which they generally perceive to have been impersonal and quantitative. There is a danger that their demand for attention to group process neglects an orientation to content material and development of specific intellectual skills such as mathematics, logic, and even effective communication through use of the English language.

2. In an urban age the cloister is commonly decried as irrelevant. While we consciously sought to guard against it, a small, innovative educational community runs the risk of institutionalizing an esoteric program and ethos which is, in effect, a cloister. The world "out there" neither understands nor appreciates the unique and "beautiful thing" going on behind the walls. A kind of paranoid life-style may result along with what is essentially a moratorium from larger social involvement and responsibility.

3. The result may in fact be a devolution of a regressive environment. The "walls" of the cloister (maintained by a staff who unwittingly become

parent surrogates) shield the members from the fiery darts of criticism by outsiders. In this kind of environment, behavior may be tolerated and even condoned that would never be functional in the larger community. It is not simply a matter of long hair and bare feet (these are superficial) but may be more serious matters of dependency, indulgence in careless or sloppy intellectual standards all in the name of creativity. Both our casual observation and our more systematic evaluation have led us to conclude that a cluster college is particularly vulnerable to faculty members who need to collect disciples around them and to students who are seeking a "guru-type" master.

4. A naive, idealistic view of the community opens the door to a kind of ethos of conformity which prohibits pluralism and the enriching quality of variety, whether in ideas or life-style. If students or faculty become so "at home" with their environment, they are rendered potentially irrelevant to a complex and struggling world. Deliberate attempts have been made through the recruitment of faculty and their assignment to teaching teams to provide for a diversity of intellectual positions rather than unanimity.

5. The college must avoid becoming a therapeutic environment per se. Emphasis upon affect may become healthy, but this should not be the goal of an educational institution. The small, innovative educational community may attract a disproportionate share of disaffected or even emotionally deprived students. It may be perceived as a haven from some of the harsher realities of a world in the last decades of the 20th century. Any small college stands in danger of becoming a sheltered workshop. This fact became apparent when a cluster analysis was made of the freshman test data on the charter class.[4]

A six-factor solution evolved and in selection of the names for each factor it became clear that several implied concepts of disability. Eventually they were labeled (1) Anomie, (2) Fatalism, and (3) Intropunitiveness.

Opportunities

The setting forth of problems such as these above may tend to obscure the inherent positive opportunities in an institution of the character of Callison College. Perhaps this is the crux of the evaluation problem: Is it possible to show a balance sheet favoring opportunities rather than liabilities in such a program?

---

[4] Ibid.

1. The greatest opportunity may rest in the solution of the problems themselves. It appears possible to utilize the tension between the affective and cognitive styles in a creative manner. In order to do this the faculty (and indeed the students as well) must be aware of the potential dangers. To be able to deal with these in a positive manner demands an articulation of goals and philosophy which enables the juxtaposition of positive and negative elements.
2. Education means change. Students come to college to experience change within themselves. Further, the institution must be open to change within itself. Education should mean the possibility of discovery and elaborating the potential of both the individual student and the educational environment. An overall question arises: Is Callison dedicated to the institutionalization of innovation and change *or* is its objective the production of certain kinds of end products, regardless of whether or not the program is stabilized or innovative?
3. The application of pragmatic questions appears to be an essentially positive opportunity. This process obviates the propensity toward the cloister. Life can be both "now" and "out there." The changing student in an innovative environment has the possibility of testing the relevance of his education. Abstractions and conceptualizations must be tested against reality. This may be subjective and affective ("Do I perceive a sense of satisfaction in what I am doing?") and objective ("Am I discovering a functional value in my contribution to the larger society?"). The present mood of student activism may be a crucial test of the place of youth in this culture.
4. Both students and faculty have opportunity to participate in the experience of evaluation without meritocratic implications or a judgmental response to evaluation.[5] The evaluation process itself could provide a social model.[6]

---

[5] "Meritocracy is a society in which social position and economic position are gained on the basis of merit. There is complete equality of opportunity; positions are allocated on the basis of talent and performance rather than inheritance or social advantage . . . the educational system is the basic institution that attempts to sort and select personnel. . . . The student in this system is terribly afraid of making mistakes; his fear permeates his relationship with the new, terribly distant figures of judgment and hell-fire, the faculty" (Gusfield, 1963, p. 6).

[6] Paul L. Dressel, present volume, p. 1.

5. Small, innovative colleges in contemporary western society may address themselves to the age-old questions regarding the value of liberal education. The affluence and technology of this age produce the first real opportunity for validating the goals of liberal education for the "masses," the typical students, as opposed to the "classes," the select elite for whom a vocation was not an economic necessity. A small liberal arts college allows for experimenting with the dynamic relationships between creative uses of leisure, the quality of life in the 21st century and the general utility of a liberal education. Callison may enrich this experimentation through its inclusion of nonwestern culture in its general education. In this process the values, affluence, and belief in progress of the West are juxtaposed with the underdeveloped economies of the East, their ancient traditions and their value sytems.

6. All these problems and opportunities provide a setting in which real educational experimentation may occur. Little genuine experimentation is taking place in education. Most innovation is simply a mechanical tinkering, a patching of existing systems and traditions.[7]

Projections for the Second Phase of Callison

The college must embrace and examine the problems and opportunities confronting it. An honest and serious reflection on these dynamics gives opportunity to project into the future.

1. The college must refine its objectives and goals. In the second phase, the objectives of the college should become more specific based upon the lessons of our history.
2. There must be an increasing involvement of the faculty in *formal* evaluation in order to protect against the separation of the founding fathers from new faculty.
3. Inasmuch as traditional criteria for evaluation are outdated for contemporary education, especially in a situation such as Callison's, new criteria must be evolved. Methodology is a critical factor. In all candor, some kinds of evaluation, especially those designs dependent upon use of a control group, are probably impossible (Feldman & Newcomb, 1969, pp. 106-145).

---

[7] Ibid.

Simply comparing Callison students to students in a traditional college actually reveals little. As a matter of fact, many changes in college students may simply be the result of time and developmental processes common to youth whether in college or not. When technical, vocational skills are removed as a goal of education, evaluation by traditional criteria are of limited value.

Attention must be given, however, to the following directions which are of a peculiarly pragmatic nature.

1. The visibility of the institution in maintaining a positive image may be a legitimate criterion. If both faculty and students continue to be attracted to the college, then it is meeting a need.

2. Closely related but a step removed is the retention rate. If the students continue to enroll and continue to completion of the program, it may be assumed that some degree of positive evaluation may accrue. In effect, if the college stays in business it may be accorded some measure of success.

3. The question of what happens to students during and after their participation in the college continues to be a valid question in evaluating a program. However, at this point in history, the age-old standards no longer hold. The sense of satisfaction and fulfillment must become the dominant criterion. Investigating former students 10 years after completion of a degree at Callison should not be predicated upon the canons of success traditionally used in this culture. Rather than ask if and to what graduate school the person went, rather than establishing a hierarchy of professions and placing the person on this kind of scale, rather than inquiring as to gross economic worth or dollar income, perhaps it might be asked if the person is happy, has a sense of worth, a sense of personal satisfaction and fulfillment related to his collegiate experience.

Summary of Evaluation Procedures

Evaluation questions need to be directed to certain recurrent features in the life of any college. Literature on experimental colleges is rich with illustrations of the features which have characterized the life span of these colleges—other worldliness and impracticality, internecine warfare and out-group hostility, and an ostrich-like avoidance of self-criticism (Gaff, 1970; Gusfield, 1963). Awareness of these recurrent features and also the pitfalls which have doomed other experimental colleges should be the focus of much of the evaluation effort in a new college.

It is important to know which features of the college are *nondistinctive* rather than indulging in narcissistic assumption that the Callison experience is unique. With awareness of common programs and problems, the possibility of sharing and borrowing from the experiences of other colleges becomes a reality. Recognizing the college's *unique* features will permit the question, "Are these features educationally relevant?" Further, attention must be directed to the question of whether or not these unique features are serendipitous, the by-product of the Hawthorne effect almost inevitably associated with educational experiments, or are they features which by deliberate planning may be repeated for future college generations?

Continuing involvement of students and faculty in the evaluation process may lessen the possibility for deceptive rationalization of success. Miles (1964) warns that the infrequency of systematic evaluation often allows "*substantive* failure of an innovation–inability to achieve desired results" to go unnoticed or to be defended against vigorously when it is pointed out (p. 659).

References

Carpenter, M. The role of experimental colleges in American higher education. In W. H. Stickler (Ed.), *Experimental colleges: Their role in higher education.* Tallahassee, Fla.: Florida State University, 1964.

Dressel, P. L. (Ed.). *Evaluation in the basic college.* New York: Harper and Brothers, 1958.

Feldman, K. A., & Newcomb, T. M. *The impact of the college on students.* Vol. 1. San Francisco: Jossey-Bass, 1969.

Gaff, J. G., & Associates. *The cluster college.* San Francisco: Jossey-Bass, 1970.

Gumperez, E. M. *Internationalizing American higher education.* Berkeley: Center for Research and Development in Higher Education, University of California, 1970.

Gusfield, J. Beyond Berkeley: High noon on campus. *Transaction,* March/April 1963.

Gusfield, J., & Riesman, D. Innovation in higher education: Notes on students and faculty encounters in three new colleges. In H. Becker, B. Geer, D. Riesman, & R. S. Weiss (Eds.), *Institutions and the person: Essays in honor of Everett C. Hughes.* Chicago: Aldine Press, 1968.

Katz, J. (Ed.). *No time for youth: Growth and constraint in college students.* San Francisco: Jossey-Bass, 1968.

Martin, W. B. *Alternatives to irrelevance: A strategy for reform in higher education.* Nashville, Tenn.: Abingdon Press, 1968.

Mayhew, L. B. *Colleges today and tomorrow.* San Francisco: Jossey-Bass, 1968.

Meyer, S. L. The University of the Pacific and its "cluster colleges." In W. H. Stickler (Ed.), *Experimental colleges: Their role in higher education.* Tallahassee, Fla.: Florida State University, 1964.

Miles, M. B. (Ed.). *Innovation in education.* New York: Bureau of Publications, Teachers College, Columbia University, 1964.

Patterson, F., & Longworth, C. R. *The making of a college: Plans for a new departure in higher education.* Cambridge, Mass.: M.I.T. Press, 1966.

Sanford, N. (Ed.). *The American college.* New York: John Wiley and Sons, 1962.

Sanford, N. *Where colleges fail: A study of the student as a person.* San Francisco: Jossey-Bass, 1967.

Watson, G. Utopia and rebellion: The new college. In M. B. Miles (Ed.), *Innovation in education.* New York: Bureau of Publications, Teachers College, Columbia University, 1964.

Webb, E. J., Campbell, D. T., Schwartz, R. D., & Sechrest, L. *Unobtrusive measures: Nonreactive research in the social sciences.* Chicago: Rand McNally and Co., 1966.

Chapter 8

NEW COLLEGE
SARASOTA, FLORIDA

John Elmendorf

There is probably no single trap into which the administrator falls more often—be he manager of a major corporation or president of a college or university—than the illusion that his problems are unique. The more his enterprise partakes of the experimental mode, the more this is likely to be true. I shall try, in these brief notes, to stay out of that trap, but I must warn my readers that I am ever susceptible to falling into it.

We are concerned in this series of essays with the "problem" of evaluation. The marketplace can solve some of these problems in the industrial sphere because evaluation, in the long run, is determined by whether or not our product is purchased by enough people, at a high enough price and a low enough cost to produce a profit. Thus, the business world, as usual, has empirical evidence to support its systems of evaluation.

It might be useful to make a few tentative remarks directed towards evaluation of what one could call the final product, *not* the process by which it was produced, a far more complex and pervasive operation which will concern us in most of this paper.

If we in the academic world could "sell" our product, the graduating senior, all might be well. We can't. We can only count fellowships, graduate admissions, job placements, or indulge in fanciful speculations on the nature of "success," and, at best, we can arrive at a value for our final product based on what may well be the least accurate criteria ever conceived. Such measures as life-style, creativity, the liberated spirit, etc., serve more to confuse than illuminate the problem.

So what *do* we do? What, specifically, does the experimental, liberal arts college do, when it is unrelated to Church or State? The "if" is important here, because the church-related institution can point to X members of its graduating class who effectively entered the life of the supporting sect and even the state-supported college or university can point to the economic impact of its graduates on the vitality of the political unit of which it was a part and ascribe the economic "good health" of its suzerainty to the input of the college graduates which it has fostered and sustained.

For the private and independent college, however, there is a different set of problems. It may, as do some, aspire to rather limited objectives, preprofessional or preoccupational, in which case again some reasonable assessments can be made on the basis of the "productivity" of its graduates. But the real question emerges—with horrifying clarity—in the college which sets for itself objectives which lie at the highest level of abstraction. New College is such an institution, and the evaluation problems of an institution such as New College, while by no means unique, are among the most arcane in the business.

There are many reasons for the emergence of these problems. First among them is the challenge posed by the college's own statement of purpose or intent. Many such colleges have spent hours, days, perhaps years refining a definition of purpose which has become more recondite with each revision, until its *stated* objectives sound as though every graduate would perforce be a blend of Erasmus von Rotterdam, Sophocles, and Martin Luther King.

Prudence—or a fear of a libel suit—keeps me from commenting on the credibility gap induced by some of the statements of purpose I have read in college catalogs. Honesty demands that there be a response to the plaguing problem of the almost universal discrepancies which obtain between such statements and the process and educational content which are the realities which provide the substance of the rhetoric.

What *are* the devices open to an honest college administrator who wants to know what is the standard deviation of his school from its announced aims (whether or not he intends to publicize the results!)? To answer the question honestly, one has to examine the structures involved. Here, I am disposed to use New College as a model, not only because it is familiar, but because it is very probably only one system of many which could be examined, most of which would share many of the same problems.

A stated aim of New College, as an experimental institution open to change, is "to provide an opportunity for its students to attain a fair degree of *general* knowledge as well as a reasonable high degree of competence in a particular field." Another is "to bring good students together with good teachers." Finally, the college has stated that "in the final analysis, each student must find a way to be responsible for his own education."

This is heady stuff. It is, however, not at all inconsistent with the stated aims of many colleges and universities; and it suffers the usual problems of having to be transmuted from rhetoric to process. Once it has become process, there is at least some possibility that it can become subject to evaluation.

It may be useful here to dismiss some of the obvious. Whether victim of "faddism" or "tinkering," New College, as a residential, coeducational, and independent college, has done away with grades, courses, fixed curriculum, class attendance requirements, and even formal degree requirements. It has set as its operational mode the thesis that a college is first of all a *learning* establishment, not a *teaching* establishment, where judgments will be made on the levels of learning achievement rather than on the degree to which students have succeeded in responding to "teaching." (We *do* have a faculty, and they *do* "teach," but this is viewed institutionally, as a necessary evil rather than prima facie need.)

How, then, do we evaluate? *What,* for that matter, do we evaluate?

We have discussed, in these essays, summative and formative evaluation. There has been a tendency, logically defensible in some ways, to suggest that these are "either-or" modes of evaluation. Our experience has been quite the opposite. The ingredients of evaluation at this "experimental" institution are not terribly different from those available in any institution. But the direction from which they are approached is very different. When one begins from the assumption that the important function of the college is *learning,* not teaching, quite a few preconditions of judgment are altered.

As empirical evidence of achievement, we have a variety of data, some objective, some subjective. On the one hand we have GRE scores, and comparable records of candidates in law, medicine, and business. Far more important to us, however, are productivity factors, the quality and range of senior theses, independent study projects, and created works of art or scholarship. Though these may sound like items not easily measured for quality, they are in fact susceptible to comparison. For example, both commercial and academic publishers have standards of excellence which must be met by aspiring writers. It is useful, therefore, to gauge how much of what is produced by students is deemed worthy of publication. Similarly, in the areas of science, research productivity is recognized by quite independent organizations, whose value judgments rarely consider factors other than merit. When student work is cited or when undergraduate students are invited to read papers to learned societies, we are dealing with reasonably empirical evidence. And finally, few critics are more discerning than public bodies, such as in the realm of effectiveness of participation in action programs. They too weigh the learning which has taken place and make hard judgments as to its value to our society. Again, the evidence is both objectively and comparatively evaluated. To a degree, therefore, these are the "summative" elements in our evaluation system.

New College, therefore, has a wide base of evaluation centered in the empirical evidence of learning, as determined by "outside agencies." This is never enough, for there are some levels of achievement which are not measurable by society at large, nor do they satisfy the "formative" evaluation concept which is also a central concern of the college.

Each student at New College "progresses" on the basis of demonstrated levels of competence in a wide variety of fields. Each faculty member makes *his* judgments on the student's performance in terms of his perception of the relation of the student's performance to an abstract conceptualization of "competence." That this is subject to being viewed as a capricious system must be obvious. What is *not* obvious, however, is that there exists a community of scholars at the college, all of whom share relatively few students. There is a tendency, therefore, for evaluations to relate to performance seen in a close relationship to capacity—as it is perceived by the corporate faculty—and a corresponding balance between personal predilection and corporate responsibility.

There still remains the question of collective—or administrative—value judgments about the efficacy of the entire operation. Here the concept of formative evaluations plays its major role. By examining discrepancies between the societal-objective judgments and the internal-subjective evaluations, the academic administrator has a ready basis for procedural decision making. If, for example, students are scoring unusually well on standardized tests or other recognized societal estimates of worth and, at the same time, are failing to achieve according to faculty standards, he can adopt one of two modes of reaction. He can suggest that the faculty may have set unrealistic standards or he can question the pragmatic validity of the "objective" measurements. On the other hand, those whose "in-house" performance has been judged exceptional, but whose external records seem out of harmony with faculty judgments, present the opposing thesis that substance may have outstripped quality and that the *esse* may not equal the *videri.* In either case–and this is perhaps the operational crux of the matter–the administrator must himself make judgments which affect the continuing structures of the institution.

From the time New College opened until very recently, there was a senior member of the staff–and faculty–known as "the College Examiner." His roles were several. As an expert in tests and measurements, he counseled the administration on the evaluation of tests, recommended or rejected certain testing devices, and acted as a guide to faculty members in the preparation of their examination instruments. In quite a different role, he himself administered both nationally standardized tests and devices of his own invention designed to evoke information of potential use to the College. Thus, for example, he was able to identify a number of new factors in precollegiate experience which could serve as predictors of certain kinds of behavior, which in turn could have internal utility in counseling and in the "in-house" evaluation system.

A final role played by the College Examiner was that of seeing that this very new institution was staying somewhere within a reasonable boundary line vis-a-vis its companion institutions. This involved comparative studies with other schools and colleges as well as specific efforts to determine the relative national status of New College students with respect to their peers. There is no doubt that all these processes were of value to a newly founded institution, but it is of singular interest that, within five years or so after its beginning, there appeared to be a well-defined limit to the utility of these evaluative procedures.

It may reasonably be asked why this very useful, even indispensable, tool lost its effectiveness as the College grew. It would seem more reasonable to assume that sheer growth would create statistical conditions conducive to even more accuracy and hence still greater predictability with its accompanying utility to the institution. There was, on the other hand, clear evidence that a genuinely innovative institution could impair its own ability to effect change by too much concern with comparative studies based on what came to be considered "false" norms of one kind or the other. New colleges beget new value systems which, in turn, demand new measuring devices.

At present, in its seventh year of operation, New College conducts its self-appraisal systems almost entirely on "formative" modes of evaluation. Through a somewhat complex system of student-faculty committees, contracts, and arbitration devices, evaluative judgments are arrived at, executed, challenged, and eventually recorded. In the absence of any grading system whatsoever, words abound and one might almost say that a controlled adversary system exists complete with a fairly elaborate complex of checks and balances. But the process is focused ultimately on the sense of the desirability and inevitability of change, which results in an almost permanent presence of vital forces acting on the evaluative process to insure that it is functioning in harmony with the status quo but, at the same time, that there is continuing evidence that it will and can function adequately within whatever frame of reference may be the successor to that particular status quo. Thus, self-study–in a communal sense–is the permanent state of affairs.

An illustration or two may be in order. Within the three-week period immediately preceding the writing of these notes, the College had a three-day meeting of its Trustee committee on academic policy and personnel, in which faculty, students, administration, and three outside experts participated. The basic objective was evaluation of program, process, personnel, and planning. In the same period, a group of students and faculty conducted a detailed survey of priorities for future growth, with specific reference to areas of academic enterprise which the community would like to see strengthened or introduced. Still, during this same period, two separate bills were introduced and discussed in the Student Executive Committee, both having as their primary thrust efforts to review or reexamine the efficacy of programs either already in effect or actively being contemplated. These were not isolated or exceptional events, but rather they are typical of the sustained efforts of the College to assess itself and both its broad aims and specific techniques for achieving those goals.

It seems absurd, in a way, to suggest that any reasonable institution should set up as one of its major goals–and maintain as one of its major activities–the need for permanent self-scrutiny and the corollary capacity to be able to change in accordance with the findings of such a process. That is, however, almost precisely what goes on at New College. It is to a large degree subjective evaluation. Any given item is therefore cause for concern lest it be seen as capricious, biased, or worse. The sum total, however, has about it a sense of objectively arrived at policy, sufficiently forceful to aid in encouraging change in the direction of broadly perceived needs. It can even produce–on rare occasions–equally clear mandates for no change or a delay in change until some more clearly defined "sense of the meeting" can be distilled.

I began this paper by referring to the "trap" of seeing one's problems as unique. It is generally conceded here at New College that we are somehow "different," but it is equally well recognized that we share many problems with others. What *may* be unique is our collective penchant for placing the evaluation process in a front-and-center position, and behaving institutionally as though the process itself were as important as the modes and qualities of learning that were being evaluated. It is at least possible that this is true, and, if it is true, it should certainly cease to be unique.

Chapter 9

EVALUATION FOR SANTA CRUZ

Lloyd J. Ring

The University of California, Santa Cruz–usually referred to as "Santa Cruz"–is one of 24 institutions around the country (Gaff, 1970) that has developed "cluster," "residential," "inner," or "subcolleges." Santa Cruz, like many of these new developments, has not been subjected to careful analysis and evaluation. Now, because of the growing implications for the rest of higher education, and because of the growing need for self-knowledge, Santa Cruz is exploring a procedure for evaluating its activities and programs.

This chapter considers the problems and the prospects of evaluation for Santa Cruz. The following sections will describe the resistances to evaluation while recognizing that evaluation is taking place anyway; outline three specific examples of the need for systematic evaluation; and suggest a theoretical framework for evaluation that may be appropriate for Santa Cruz.

Resistance and Evaluation

Suchman (1967) observes that few new programs ever include a plan for the evaluation of their development. Such is the case with Santa Cruz. This situation is not unexpected in the light of the problems inherent in conceptualizing, describing, and selling a new idea to a Board of Regents and a State Legislature. Although the presentation of the idea was laced with promises of success, neither the author of the idea nor the "buyer" of the idea had the desire to put it to the test too early in the game. Once the venture was underway, additional forces came into play to further delay serious criticism or evaluation. Santa Cruz took on some of the characteristics of utopia for both planners and faculty, and many of the goals, purposes, and structures of the new campus were taken on faith and with deep personal commitment. As Gaff (1970) observes, to question the utopian order–let alone to seek empirical knowledge–is tantamount to heresy. At the same time, Trow (1969) recognizes that some innovations in education are justified by their own intrinsic qualities.

Evaluation was strongly resisted by the early members of the Santa Cruz campus. There was an overwhelming concern that measurement might interfere with the free development of community and that too many guidelines and restraints would emerge. There was also a feeling on the part of some that the only evaluation necessary was the test of history. They contended that after the campus had developed, it would be possible to look back and evaluate (or rationalize) how successful the endeavor had been. Such an approach has already taken place and the first four years of Cowell College are recorded in a student-written history entitled *Solomon's House* (Cowell College Students, 1970).

Early statements about Santa Cruz were hopeful and idealistic, and implied that Santa Cruz was the antidote for the ills of the large impersonal university. Not only did the statements imply a high order of value but they also presented problems for evaluation.

> The Santa Cruz campus, which expects to grow to 27,500 students within thirty years, will seek to organize teaching in such a way that the advantages of a small college–particularly opportunity for discourse–are combined with those of a large university–great scholars, excellent libraries and laboratories, and a rich and varied cultural life. We hope to overcome the too common separation of inquiry from teaching, of one discipline from others, of faculty from students [McHenry, 1965].

Other phrases that conveyed vivid images but defied specification were: "the essence of Oxford," "the colleges will be intellectual states within a federal university," "make the university seem small even as it grows larger," "each college will develop its own intellectual center of gravity," and "each college will be devoted to the 'liberal arts.' " Even in the guidelines for curriculum, there were statements that defied operational definition: "to teach habits of intellectual honesty," "to transmit knowledge, and with it understanding of the significance, methods, and interrelations of our various ways of looking at the universe," and, "by enlarging his understanding of his own and other cultures to develop the student's ability to stand outside himself, to understand as a consequence his location and opportunities." Many of these statements have persisted through six Santa Cruz catalogs.[1]

These, then, are the statements that excited the imagination, generated expectations, attracted faculty and students, and launched a new campus. The ideas conveyed hope for a humane environment, for a closely knit community, and for a great university; but they gave few clues about how to determine whether or not the hopes were realized. Few goals were ever made specific and no procedure emerged for monitoring or collecting information on programs. Those criteria against which performance could be measured were extrinsic to the campus: the honors of the faculty attracted to Santa Cruz, the grade-point average and the SAT scores of students accepted, the number of Woodrow Wilson Fellowships awarded graduates, and the prestige of the graduate schools which accepted Santa Cruz students. Intrinsic criteria usually included self-satisfaction, the number of students known to a faculty member, the intellectual caliber of the students, and the degree of praise bestowed by duly impressed visitors to the campus.

One persistent restraint to evaluation has been that of staff turnover. Responsibility for evaluation has been vested in the Office of the Director of Academic Planning. The first Director was deeply involved in planning for the early colleges and while he recognized the need and even established a Committee for Educational Research, he never made much progress on evaluation. The second Director held the office for less than one year and the office has been removed from the staffing list for the past two years. The impact of staffing is further highlighted by the introduction of three new Vice Chancellors, one new Provost, and one new Assistant Vice Chancellor for Planning and Analysis.

An emerging acceptance for systematic evaluation or "feedback" is indicated in several explicit ways, but it will take some time to obtain understanding of and approval for the specific components of an evaluation program. The observation of Trow (1969) seems applicable to the Santa Cruz situation as the campus moves through its sixth year of operation: ". . . the time to assess an innovation is when it is no longer an innovation, when it has become routinized and no longer can call forth the special energies, resources and enthusiasms of an 'experiment' [p. 23]." Although no formal process has emerged, there is evidence of evaluation in all areas of the campus. It is evident in the allocation of resources–programs expand, courses are added, faculty and staff are hired, students are admitted, new colleges and facilities are planned. Few changes have been made in the original long-range development plan or in the academic plan. Decisions are made continuously on implicit or explicit values, judgments, or evaluations. Dressel (1961) states the process succinctly: "When one is faced with choice, evaluation, whether conscious or not, is present [p. 6]." The lack of a systematic process of evaluation, adds Dressel, permits "prejudice," "tradition," or "rationalization" to weigh heavily in decision making. Santa Cruz has arrived at the precipice in this respect for as new colleges are planned, additional student housing contemplated, new academic programs introduced, and student growth and development considered, there is little but original convictions, "traditions," and "rationalization" to serve as guidelines for decisions.

---

[1] Most of these statements appear in UCSC catalogs, many beginning with the opening year, 1965-66.

Although Santa Cruz lacks a plan for formal evaluation, many individual projects have been undertaken. In some cases, the information acquired has influenced decisions in specific areas. In other cases the information has been sought out of curiosity and has found a resting place in a dusty file cabinet. In a few cases, noncampus agencies have collected and retained information for inclusion in publications of their agency. A partial list of these projects follows:

Sponsored by faculty and/or students:

1. Ethnic Studies Committee–survey of students regarding plans for College Seven.

2. Offcampus Housing Survey–to determine attitudes, values, and quality of community housing for students.

Sponsored by colleges:

1. *Solomon's House*–a self-conscious history of Cowell College.

2. Student Attrition–a study of Merrill College attrition.

Sponsored by a campus-wide agency:

1. Academic Plan Revision (1970-80)–broad participation in reviewing all aspects of the academic plan.

2. Faculty Course Load Survey–divisional and campus-wide survey of courses taught and students enrolled by Board of Studies, Divisions, and Colleges.

3. Student Attrition at UCSC–survey of student attrition from 1965 through winter quarter 1967.

4. Survey of UCSC Graduates–survey of the graduating seniors of 1968, 1969, 1970.

Sponsored by Office of the President:

1. Faculty Effort and Output Study–a nine-campus survey of faculty effort and output.

Sponsored by various noncampus agencies:

1. Accreditation Self-Study–a self-study report for and evaluation by the Western Association of Schools and Colleges.

2. Center for Research and Development in Higher Education, Berkeley.
    - a) Gaff and Wilson–faculty participation for a study of faculty cultures.
    - b) Heist–longitudinal study of students (see Heist and Bilorusky 1970 for partial report).
    - c) Martin–a study of eight institutions which included Santa Cruz (Martin, 1969).

3. University of Michigan–"An Inter-University Program of Research on Residentially Based Units for Undergraduate Study"–a five-institution consortium study.

This partial list of projects represents a wide range of research and evaluation concerns. If the campus had been prepared to take advantage of these and other research opportunities through an office of evaluative research, considerable data would already be available. Coordination becomes more crucial each year as students, faculty, and noncampus agencies request or simply undertake surveys and studies for their own particular purposes.

Are Promises Made, Promises Kept

To illustrate the growing need for a systematic evaluation of all aspects of Santa Cruz, three specific topics (goals and purposes, students entering-students leaving, and student housing) have been chosen for analysis. These topics are crucial to the campus and the analysis will indicate changes

TABLE 1

Academic Planning Statements of Goals and Purposes Santa Cruz Campus, 1962 through 1970

Statements Implying More than Casual Emphasis		1962	1965	1970	Cowell 1964
1. Campus will grow by colleges	*	yes	yes	yes	yes
2. Colleges will be the principal source of innovation	*	yes	yes	yes	yes
3. Campus will stimulate academic innovation	*	yes	yes	yes	yes
4. Small class size will be a campus goal		yes	yes	yes	yes
5. High priority for quality graduate programs		yes	yes	yes	
6. Academic standards will be of high quality	*	yes	yes		yes
7. Colleges will emphasize interdisciplinary work	*	yes	yes		yes
8. Campus will emphasize student writing skills		yes	yes		yes
9. Campus will foster personal growth of students	*	yes	yes		yes
10. Colleges will have disciplinary themes		yes	yes	no	yes
11. Campus will maintain "breadth" requirements	*	yes	yes		yes
12. Administrative plans for academic change included		yes	yes	no	yes
13. The proliferation of courses will be avoided		yes	yes		yes

14. Teaching aids and technology will be urged		yes	yes		yes
15. Residence facilities will support academic program	*	yes	yes		
16. Degree attainment will be accelerated by examinations		yes	yes		
17. Faculty research has a high priority	*			yes	
18. Colleges will be designed on several models	*			yes	
19. Size of campus may be less than 27,500				yes	
20. Colleges will develop core programs					yes
21. Colleges will develop cultural programs					yes

*These items were also included in the recent survey.

Note.–Of the 16 items included (marked "yes") in both the 1962 and the 1965 documents, only 5 appeared in 1970.

over a period of years. Evaluation is implicit in the analyses, but they also illustrate the need for a more systematic approach to evaluation.

1. Goals and Purposes

Santa Cruz is the product of needs observed and the high ideals held by a few individuals. These were shared and infused with the hopes and energies of others; were altered and compromised by the ever-expanding community of students, faculty, and planners and by the constraints and fickle fortunes of the larger University system. To obtain some insight into how the goals and purposes for Santa Cruz have fared over the past eight years, an analysis of four academic planning documents was undertaken, and an interview survey was conducted with a selected sample of the campus community to check personal preferences against planning statements. This project represents a first step in the process of evaluation that may stimulate the development of continuous evaluation.

This analysis of goals and objectives does not pretend to be exhaustive.[2] It is based on selected statements found in four planning documents:

1. *A Provisional Academic Plan for the Santa Cruz Campus, 1965-1975.*
2. *Cowell College: A Statement of Aims, 1964.*
3. *Santa Cruz Campus Academic Plan, 1965-1975.*
4. *Proposed Academic Plan for the University of California at Santa Cruz, 1970-80.*

and in the *University of California, Santa Cruz Catalog,* 1965-66 through 1970-71.

Two of the missing statements reflect deliberate changes in goals: (10) College themes have become a subject of debate with disciplinary themes losing favor. (12) The role of administration was perceived to be the facilitation of change rather than the planning of change. The other nine items reflect goals that have been "taken for granted" or that have become less salient. Possible explanations are offered for the following: (11) The language requirement has been dropped this year and other breadth requirements are under review. (13) The proliferation of courses has become a reality. (14) Instructional technology has been resisted. (16) Although the option is available, students have not taken advantage of it.

---

[2] The author is indebted to associate Robert Bosler, Administrative Analyst, for much of the analysis and survey materials in this section.

There are three new items in the 1970 list that are of some significance. (17) While the structure of Santa Cruz is radically different than other campuses of the University of California, the reward system is only slightly altered. The ambivalence attached to university standards and campus performance prompted the inclusion of this item in 1970.[3] (18) Evaluation of the first five colleges, recognition of possible budget restrictions in the future, and a desire to innovate further supported speculation about different models for colleges. (19) Internal resistances, external community pressure, and a genuine desire to seek an optimal size for a campus like Santa Cruz have aroused the need for careful study of the previously proposed ceiling of 27,500 students.

The items drawn from the Cowell College statement of aims provide a comparison of the goals of a single college with the goals for the campus. The differences observed relate to concerns primarily relevant to the liberal arts colleges *vs.* the university. Cowell College is the only college that developed such an elaborate statement of aims prior to its opening; being the first one, its statements naturally reflected those for the University.

Statements of intent do not guarantee results or even agreement, so a survey was devised to check some of the statements against the personal perceptions of a selected sample of the Santa Cruz community. The survey sample included the chancellor, four of the five vice chancellors, all six provosts, five faculty members, six undergraduates, and two graduate students. Faculty members were from four of the operating colleges and from each of the three academic ranks. Students represented each college and had been at Santa Cruz two or more years. A letter, a check list, and an interview guide were sent to each of the participants in advance of a personal interview. The survey/interview had five distinct, interrelated parts. The purpose of each and the results obtained follow:

Part I. Seventeen phrases representing possible goals for the Santa Cruz campus were listed and participants were asked to check those they thought should have high priority in the next five years. Table 2 lists the results.

---

[3] The problems of a change in structure without a change in the reward system will be dealt with in a forthcoming paper by two campus economists, Robert F. Adams and Jacob B. Michaelson.

TABLE 2

Responses to Preinterview Questionnaire, by Number Checking Each Statement, and Listed in Descending Order by Totals

Possible High Priority Goals		Vice Chancellors*	Provosts	Faculty	Students	Total
1. Greater emphasis on interdisciplinary learning and research	**	5	4	4	5	18
2. Greater support and stimulation for academic innovation	**	4	3	4	6	17
3. Increased college responsibility for general education and/or academic innovation	**	5	4	4	4	17
4. Encouragement of college individuality		5	5	2	4	16
5. Thoughtful evaluation of academic quality	**	4	4	2	6	16
6. Encouragement of better student/faculty/administrative relationships		2	3	2	6	13
7. Greater integration of the residential houses with the academic program	**	3	2	2	5	12
8. Creation of new models for colleges	**	2	3	2	5	12
9. Careful study of the short-run and long-run effects of evaluation grading		1	3	3	4	11

10. Greater interest in fostering the personal growth of students	**	2	4	0	5	11
11. Continuing campus growth by colleges	**	3	3	2	2	10
12. Greater emphasis on student enrollment patterns in the allocation of faculty		2	0	3	2	7
13. Reduction of the number of course requirements	**	1	1	1	3	6
14. Reduction of administrative and committee tasks for faculty		2	2	0	2	6
15. Colleges setting the criteria for disciplinary majors		0	2	0	3	5
16. Higher priority for faculty research	**	0	1	0	1	2
Persons Interviewed		5	6	5	8	24

*Includes the responses of the Chancellor.
**Based on statements listed in Table 1.

The data of Table 2 are extremely tentative for, although the survey sample was selected to get a cross-section of the campus, it would not necessarily be representative of the total campus community. Moreover, the translation of some of the goal statements into survey items is less than precise. However, there are enough questions generated by the data to warrant a large-scale survey of this type with appropriate survey technology.

The most significant aspect of the survey was the reaction of the participants to the study. Both interviewers were impressed with the receptivity and depth of interest expressed by each participant. It was evident that Santa Cruz was a significant part of their lives, and they had both criticisms and hopes for its future.

The 10 items based on goal statements from Table 1 are interspersed in the rank ordering among more recent priorities. These items represent interests and needs presently appearing in conversations around the campus. Item 4 (college individuality) reflects the increasing decentralization that is giving the colleges greater autonomy. Item 6 (better student/faculty/administration relationships) is generating sympathetic support as the campus confronts possible threats from outside the campus. Almost half the respondents recommended careful study of the effects of evaluation grading, item 9. Responses indicate that the present system is greatly preferred, but there are improvements needed and dangers to avoid. Item 15 shows the beginning of a growing mood that the colleges should be much more responsible for undergraduate programs. Note that the Provost and students are the only respondents to this item.

The table raises several important observations that indicate the need for further study. Item 10 shows that almost half of the respondents checked a concern for the personal growth of students but not one of the faculty members did so. If this observation holds for a large proportion of the faculty, much that is important to students would be diminished or lost. This observation could mean any of the following: (1) personal growth is not considered part of the academic experience, (2) faculty are ill prepared to cope with such concerns, (3) faculty perception of personal growth means being available and friendly with students, or (4) it does not have as high a priority as the other items checked by faculty. Item 12 has come to the surface as resources have become scarcer and as student enrollment patterns have changed. Those disciplines and divisions which attract increasing numbers of students are demanding priority in the allocation of faculty. The Provosts, on the other hand, see this as a potential threat to the equilibrium of the desired college mix of disciplines. No faculty member expressed the importance of having administrative and committee tasks reduced as suggested in item 14. Item 16 undoubtedly reflects a predominant opinion at Santa Cruz, but it also raises a very significant problem–what is the appropriate role of research at Santa Cruz and what does the future hold for faculty if the reward system is not supportive of the preferred role?

Some of the foregoing observations may have been influenced by the response set of the participants. The range of items checked and the median number of items checked are illustrative (see Table 3).

The faculty responses were much more constrained than the other groups and either suggest an intensity for certain goals or a much narrower interest in the possible range of goals.

Part II. Six of the items included in the Preinterview Questionnaire were arbitrarily selected for further study in the interview. The six items were chosen because they seemed to represent a range of concerns and interests. The items were printed on self-adhesive labels and the participants were asked to affix the labels to a sheet of paper in order of their personal preferences. These decisions were made by the participants before the interview began.

TABLE 3

Range of Responses and Median Number of Responses to the Preinterview Questionnaire

Respondents	Range	Median
Vice Chancellors	5-12	9
Provosts	5-13	8
Faculty	5-6	6
Students	4-15	7
Total	4-15	7

The Kendall Coefficient of Concordance–"W"–was applied to the total sample and to each of the participant categories.[4] The "W"s for each category were: *Total*, .475 (significant at the .001 level); *Vice Chancellors*, .405 (not significant); *Provosts*, .820 (significant at the .01 level); *Faculty*, .309 (not significant); and *Students*, .674 (significant at the .01 level). Siegel (1956) indicates that a high or significant "W" may be interpreted to mean that the observers are applying essentially the same standards in their rankings. Based on the present data, it might be inferred that Vice Chancellors (which includes the Chancellor) and Faculty, as groups, were not applying the same standards or values when they ranked these six goals. There is no immediate explanation for the faculty situation, but one possible explanation for the lack of a common standard among the Vice Chancellors is that two of them are new to Santa Cruz with just a few months' exposure to the campus. This possible explanation must not be construed as detrimental for it has yet to be determined whether the "established" view o[illegible]tandard is the more appropriate. The best estimate of the "true" rankings for the three categories, where "W" is significant, is presented in Table 4 for comparison.

TABLE 4

Best Estimate of "True" Ranking of Six Goals by Provosts, Students, and Total Sample

Category of Respondents	(A) Academic Innovation	(B) College Responsibility	(C) Interdisciplinary Learning	(D) Evaluation Grading	(E) Residential Programs	(F) Faculty Research
Provosts	2	1	3	5	4	6
Students	1	3	2	5	4	6
Total Sample	1	2	3	5	4	6

*Complete items are respectively 2, 3, 1, 9, 7, 16 of Table 2.

---

[4] Following Siegel (1956).

Reference to Table 2 confirms a rather congruent pattern between the ordering of these six items under free choice and their ordering under specific instructions. Items 2, 3, and 1 of Table 2 were most frequently checked and their ranking is very compatible with the close ranking of items A, B, and C in Table 4. Items 7, 9, and 16 are in descending order in Table 2 and are congruent with items E, D, and F in Table 4.

Part III. Students, faculty, and staff are attracted to Santa Cruz primarily through the process of self-selection. Some aspects of the self-selection process could be related to the stated goals and purposes of the campus. The first question on the interview schedule asked, "Why did you decide to come to Santa Cruz?"

Responses to this question suggest that individuals are less attracted to Santa Cruz' stated goals than by generalized impressions and personal expectations. Several individuals admitted being quite disappointed in their first year because of a disparity between what had been anticipated and what was experienced. Most of the responses seem to fit under six general headings: personal, academic, colleges, setting, professional, and newness.

Personal: Responses from vice chancellors, provosts and faculty comprise this category. Almost half of the responses related to discontent at previous institutions, several referred to existing relationships with people at UCSC or to anticipated relationships, several mentioned money and prestige, and a couple indicated a long-standing interest in the University of California.

Academic: Reasons under this heading were varied and often vague. Included were: mention of academic innovation, pass/fail grading, undergraduate teaching, less conflict between research and teaching, first-rate faculty, graduate work in own field, liberal arts, academic freedom (student), and the desire to be in a university. These general responses came from each of the four survey categories.

Colleges: Enthusiasm for the colleges was given as an attraction to Santa Cruz by 13 individuals who were almost equally divided among vice chancellors, provosts, faculty, and students. Five more individuals cited either the "human scale" or the student orientation of the small units.

Setting: The physical setting of the campus was the unabashed reason given by eight individuals—half of whom were students. Four more people expanded on this theme and included the whole Santa Cruz area.

Professional: This is another heading related to faculty and staff responses. For four of the vice chancellors and provosts, Santa Cruz was a professional challenge either in the creation of something or in the management of a project that is past the "pioneer" stage. Two faculty members felt their job offers were the best available, and one faculty member looked forward to close collaboration with associates in a variety of fields.

Newness: There is an intrinsic value in newness for at least five individuals.

It is evident that there is little direct relationship between personal attraction for Santa Cruz and the specific statements found in planning documents and catalogs. It was also evident in the interviews that there is great interest in these goals, and the opportunity to think about them and to discuss them would be advantageous for both the campus and for the members of its community.

Part IV. Each of the six selected goal statements was pursued in some depth during the interviews. The statement was related as a continuing goal of the campus and opinions of success or failure were solicited. The general impressions are clear.

All categories of respondents were critical but simultaneously optimistic. Innovation was acknowledged in the founding of the campus, but few examples of academic innovation since that time were cited. Colleges are seen as more innovative than boards of studies, but there are doubts about the quality of most of their programs. There were few positive comments about the creative use of residential living as part of the academic program. Students, particularly, thought much more could be done in this area and were perhaps the most enthusiastic of the respondents. Research, in the traditional sense, did not fare well. It seemed to be in need of a new definition if it was to be in tune with other aspects of Santa Cruz. Work done researching course materials, working in other fields to make interdisciplinary activities effective, and work on a college core course require the same thoroughness as research in most disciplines yet the outcome has no publishable market. If research is to be confined to such conventional aspects as laboratory work and publishing, Santa Cruz will have a difficult time achieving its goals for the colleges.

Interdisciplinary work had its share of critics as well. Those few efforts deemed truly interdisciplinary have required a particular breed of faculty or a particularly appropriate subject area such as environmental studies. Pass/fail (P/F) evaluation grading received near unanimous approval, but there are severe criticisms of certain aspects. Faculty and staff are emotionally committed to it but are not at all sure whether it is beneficial or harmful to student development. Faculty report less "nitpicking" on the part of students over grades. Students, while praising the system, readily admit that the written evaluation accompanying the P/F produces anxiety. There are also student complaints about the quality, completeness, and depth of the evaluations.

Part V. Every system develops roadblocks or constraints to the achievement of its goals and purposes. The final part of the interview asked participants to identify constraints in the system. Participants were also asked to suggest topics for study about Santa Cruz. Many of the suggestions in both categories that follow are examples of needed study, discussion, and evaluation.

Perceived Constraints

- confusion between boards of studies and colleges
- ambivalence of faculty–guardians of standards or promoters of innovation?
- personality and administrative style of various decision makers
- "publish or perish" heritage
- ineffectiveness of Academic Senate
- poor communication among students/faculty/staff
- 40-60 lower division-upper division aspect of the master plan
- proliferation of central units like the Natural Science and Social Science buildings
- programs like mathematics honors program which restrict interdisciplinary work
- science faculty resistance to participation in college life
- lack of college autonomy
- the academic "smorgasbord"

- too much student involvement in peer groups
- faculty who "prostitute" themselves to student values.

Two statements are quoted but the authors will remain anonymous:

- "so much pluralism, so much freedom, so little humility"
- UCSC is "trying to seem large while being small."

Suggested Topics for Study

- pass/fail evaluation grading
- optimal size for UCSC
- what happens to UCSC graduates?
- academic advising system
- campus housing problems
- need for increased counseling services
- how many faculty are absorbed in the value system of the discipline?–how many are seeking an alternative?
- teaching loads of faculty
- perceptions of the power structure
- police and security procedures
- use of technology in instruction
- drug use and abuse
- what can we do about graduate programs?
- use and abuse of independent studies

- evaluation of teaching for merit and promotion
- statistical analysis of college enrollments, dropouts, and transfers
- impact of the campus Educational Opportunities Program
- rationale for distribution of resources.

Because of this brief survey of 24 members of the Santa Cruz community, the number of interested, concerned, and involved citizens has increased. The information obtained will serve as a base for a wider, more carefully designed survey and many of the suggestions will stimulate specific studies of important issues.

2. Students Entering–Students Leaving

Differences among students entering Santa Cruz and other campuses of the University of California have not been carefully studied nor has there been an examination of the differences among students choosing the various colleges at Santa Cruz. The descriptions which follow are based on very primitive data and analyses, but there are indications of differences that encourage the development of appropriate information procurement processes.

Since 1966, Santa Cruz has participated in the American Council on Education (ACE) Survey of Entering Freshmen (1966, 1967, 1968, 1969). The data have not been used in a systematic manner to evaluate the changing character of the student body. Naturally, these data do not tell the whole story since a large number of junior transfer students are also admitted to Santa Cruz and there are no camparable data available for them. The data have been provided on a campus-wide basis and do not permit comparisons among colleges.

The ACE Survey of Freshmen contains more than 20 categories and hundreds of specific entries that a freshman can check to describe himself and his background. Most of the responses indicate a comparability with freshmen of the University National Norms group. There are, however, noticeable differences on a number of characteristics. Ten of these characteristics have been selected to illustrate the differences and changes in the entering Santa Cruz freshmen over a four-year period. Table 5 contains these comparisons.

TABLE 5

Selected Characteristics of Santa Cruz Entering Freshmen Compared with University National Norms for 1966 through 1969 (From ACE Survey)

Selected Characteristics	1966		1967		1968		1969	
	*SC**	*N***	*SC*	*N*	*SC*	*N*	*SC*	*N*
Average grade in high school, A+ A A–	59.7	21.7	63.6	23.8	56.2	20.8	53.6	29.9
Member of scholastic honor society	86.5	36.6	85.9	39.9	76.7	36.5	73.7	35.6
Father's education–college or postgraduate degree	60.4	32.5	64.2	36.0	65.0	33.6	64.1	35.4
Mother's education–college or postgraduate degree	44.2	20.6	57.1	23.1	50.7	21.1	45.7	22.6
Highest degree planned–PhD	35.2	12.0	36.4	14.4	35.0	13.7	33.5	14.1
Probable career occupation–undecided	12.1	4.4	30.2	10.7	32.9	12.2	37.3	12.8
Considered essential–be an authority in field	63.5	67.7	62.3	70.7	46.6	60.7	47.7	61.9
Considered essential–succeed in business	23.4	52.9	16.3	47.0	12.0	84.8	14.1	44.7
Religious preference–none	32.8	8.2	35.9	10.2	40.5	11.7	46.8	15.7
Racial background–non-Caucasian	1.9	5.0	3.5	6.4	7.6	8.4	11.5	7.2

*SC = Santa Cruz
**N = National Norms Group

The University of California has uniformly high academic standards for admission, so it is not entirely unexpected that more than half of the Santa Cruz freshmen had high scholastic grades in high school. The same is true for membership in a scholastic honor society. The ratios in both of these categories have decreased slightly as the campus sought a somewhat broader range of students.

Data on the father's education, the mother's education, and the intention to achieve a terminal degree have remained very stable over the four-year period for Santa Cruz as well as for the norms group. Indecision about a probable career occupation has shown a three-fold increase since 1966 for both populations, but for Santa Cruz this now represents more than one-third of the entering freshmen. To be an authority in a field and to succeed in one's own business have both been considered essential by a decreasing proportion of freshmen. In this case, the Santa Cruz decrease is consistent with the norms group, but it is slightly larger. The proportion of entering freshmen indicating no religious preference has increased to almost half and the percent of minority students among entering freshmen has increased each year.

It is apparent from these few observations that freshmen entering Santa Cruz are distinctive in many ways and that there are noticeable changes occurring in certain characteristics over time. Unfortunately, comparison data on other University of California campuses are not available to check some of the exaggerated differences between Santa Cruz and the University National Norms group.[5] It is also unfortunate that these data do not permit longitudinal comparisons with the same individuals. For instance, graduating seniors, June 1970, were asked to respond to two of the same categories that were on the 1966 freshman survey. The profiles for the two groups prompted certain insights and questions but valid comparisons were not possible.

The data from June 1970 graduates do permit some primitive comparisons by colleges, however.[6] No statistical tests have been performed on these data and the following observations are based only on the percent of students checking the item. The two categories were: (1) Indicate the importance to you personally of each of the following (17 items were included with possible responses of–essential, very important, somewhat important, not important). (2) Rate yourself on each of the following traits as *you really think you are* when compared with the average student of your own age (21 items were included with possible responses of–highest 10 percent, above average, average, below average, lowest 10 percent).

---

[5] **Data were obtained from the fall 1966 freshman classes on seven of the nine campuses of the University of California. Santa Cruz freshmen were higher than freshmen at the other campuses on: percent whose fathers had a college degree, percent who had A– or better in high school, percent planning on obtaining an advanced degree, percent checking above average on political liberalism, and percent checking above average on intellectual self-confidence. (Source: Office of Analytical Studies, University of California, 1967)**

[6] **Data from a survey of graduating seniors, 1970, by Robert Bosler.**

Observations on objectives deemed to be *essential* or *very important* only indicate a consistency of student response across all four colleges on four items.[7] "Helping others who are in difficulty" and "keeping up to date with political affairs" were the first and second highest responses in each college. At the other end, "becoming an expert in business and finance" and "becoming an outstanding athlete" were, respectively, least and second least frequently chosen across all colleges.

Compared to the mean response from all graduates in the survey, responses from some colleges were considerably higher than the other three colleges. Cowell College had a larger proportion of graduates indicating that it was essential or very important to: "become an expert in finance and commerce," "have administrative responsibility for the work of others," "create artistic works," and "become an outstanding athlete."

Crown College had the largest percent of graduates who checked the importance of "obtaining recognition from my colleagues for contributions in my special field" and "never being obligated to people." Merrill College led all other colleges on the importance of "achieving in a performing art." Stevenson College did not have any item more prominent than the other colleges.

Colleges did seem to group into pairs on some of these items. Cowell College and Merrill College were both considerably higher than Stevenson College and Crown College on the importance of "becoming an accomplished musician," "becoming a community leader," and "keeping up to date with political affairs." Stevenson and Crown showed higher percentages on the importance of "making a theoretical contribution to science." Crown and Merrill were tops on "writing original works." Cowell and Crown graduates gave greater importance to "being successful in a business of my own" than did those of Stevenson and Merrill.

---

[7] College V had only 11 graduates in 1970, and they are not included in these observations.

Observations on how the graduating seniors from the four colleges perceived themselves on a variety of traits show a similar pattern. Considering only the responses that were in the *highest 10 percent* and *above average* categories, "academic ability" was checked by the largest percent of seniors in each college. By percent of seniors checking, three of the four colleges showed "political liberalism" and "understanding of others" to be second and third highest. The ranking for Merrill College on these two items was reversed with the second largest percent checking "understanding of others." Each college showed the lowest percent of seniors describing themselves in terms of "political conservatism."

Differences among the responses of students by colleges showed Cowell College seniors to rate themselves highest on "social self-confidence." Stevenson College students were highest on "popularity with the opposite sex." Crown College seniors rated themselves highest on "athletic ability," "mechanical ability," and "stubbornness." Merrill College students produced the highest self-ratings on "understanding of others."

Responses from seniors seemed to pair by colleges on at least three traits. Cowell and Merrill students rated themselves higher on "artistic ability" and "general popularity." Students from Cowell and Stevenson clustered higher on "cheerfulness" than did Crown and Merrill students. Cowell and Crown students in greater numbers rated themselves above Stevenson and Merrill students on "leadership ability."

The data from the ACE Survey of Entering Freshmen and the data derived from the use of selected categories of the Survey with graduating seniors illustrate that Santa Cruz students are different than other students on certain characteristics. Moreover, Santa Cruz students show differences by colleges on the same campus. It is essential that reliable information of this nature be collected and evaluated systematically and continuously. Only in this way will it be possible to accurately portray the campus and its colleges to potential students.

The extremely complex process of selecting, self-selecting, and settling for alternatives must be examined. At Santa Cruz it can best be done by a carefully designed process of determining the characteristics of students who apply to Santa Cruz, of tracking accepted students over their stay on the campus to determine patterns of activity including courses, programs, peer group involvement, cultural and social interests, and to compare the graduating senior with his own characteristics–personal, social, academic–as he departs from Santa Cruz. This basic information on a particular student would then permit periodic checks over any number of years to determine, if any, the long-run impact of Santa Cruz.

This brief discussion has concentrated on examining one aspect of the campus. There are many aspects that need to be identified and monitored. One example would be the impact of the college/campus on junior college transfers who have not had the two earlier years' experience. How do they compare themselves to continuing juniors? What impact will the increasing ratio of upper division to lower division students have on the colleges?

There are many aspects of the programs of the colleges that have an impact on the student. Some of the specific activities that require evaluation are the following: the academic advising program; college programs (core courses or other courses); the impact of pass/fail evaluation grading; the role of disciplinary, interdisciplinary, college, and individual majors; class size and closeness of instruction; the effects of independent study; the reasons for attrition; and many others.

3. Student Housing–A Problem Unveiled

There is certainly no doubt that the University of California needed an antidote for the impersonality inflicted on students at Berkeley and Los Angeles by virtue of their size and their devotion to graduate studies. It should not be surprising then that two of the three new campuses planned for opening in 1965 would emphasize models designed to offset the impersonality of largeness. Only the Santa Cruz campus, however, placed a major emphasis on undergraduate students in the early years of its development.

The Santa Cruz model was unique in public higher education because its plan was to grow by adding small, liberal arts colleges, each containing 30 to 50 faculty members, 500 to 800 undergraduate students, and enough staff and facilities to provide as much autonomy as possible. The colleges were to be *residential* units to accommodate two-thirds of the undergraduate enrollment. This commitment to residential facilities implied more than a casual interest in an environment that integrated the living arrangements with the learning process. With five colleges built and two more in the planning stages, with tighter budgets ahead, and with evidence of decreasing occupancy, the campus has begun to reexamine the residential aspect of the colleges.

A review of the planning documents discussed earlier gives the impression that the role and purpose of the residences were overlooked along the way. The explicit description of a "residential college" in the Provisional Academic Plan of 1962 implied that the whole college, including the residence facilities, was to be an educational unit that would combine the academic and cocurricular aspects of undergraduate life. The 1965 version of the Academic Plan expanded on the earlier statement to include a description of the intellectual and more adult atmosphere anticipated in the residence halls because of the proximity and shared interests of faculty and students. The Proposed Academic Plan for 1970-80 contains only passing reference to the residential houses although recognition is given to the problems and prospects that have yet to be explored.

A similar pattern appears in the campus catalogs. In 1965-66 and 1966-67, the catalog contained a section titled "The Undergraduate College." This same section was titled "The Residential College" in the 1967-68 and 1968-69 catalogs. The section in the last two catalogs is simply titled "The Colleges." Since 1967-68, the colleges themselves have been providing the copy for college descriptions and the decreasing emphasis on the residence facilities as part of the college experience culminates in the 1970-71 catalog. There is no mention of the residence facilities in two of the descriptions, there is a brief note on residence facilities (i.e., housing) in two descriptions, and only one description includes a rationale for the inclusion of the residence in the academic life of the colleges.

Analyses of planning documents and campus catalogs are post hoc and provide no more than one indication that a problem has been slowly building over a period of years. However, when campus administrators were confronted with the following facts, the residences were perceived as an *immediate* problem.

1. The residential housing operation incurred a substantial deficit in 1969-70.
2. Statistics showed a clear decline in occupancy rates for several years.
3. Observation of the decreasing occupancy level by University-wide planners threatened the development of residential facilities in future colleges.
4. Student life-styles have changed all across the country and the double- and single-room facilities available in university residence halls are less attractive to a growing number of students.

5. There is a dearth of systematically collected information that is available for analysis by decision makers.

At Santa Cruz, analysis of the information that was available has revealed several possible sources of the problem.

1. Housing facilities have been added each year as each new college opened but with one exception the average level of occupancy has decreased each year. The following occupancy levels are given as a percent of the total spaces available and filled per academic year: 1965-66, 98.8 percent; 1966-67, 100.0 percent; 1967-68, 96.9 percent; 1968-69, 95.9 percent; 1969-70, 86.4 percent. The first financial deficit was experienced in 1969-70 when the occupancy level was considerably below the anticipated level of 96.0 percent. The following discussion examines some of the more obvious reasons for the decline.

   Housing at Santa Cruz has been optional for students since the campus opened, so the element of voluntary choice to live on or off campus has been primarily constrained by parents and the availability of offcampus housing. Both of these factors have been consistent–particularly the latter which remains at a very low level. As the campus has grown, however, there has been a drastic change in two interdependent patterns–occupancy and enrollment. Tables 6 and 7 highlight this situation.

TABLE 6

Percent of Average Enrollment Living in Residence Halls, by Level for Selected Years

Year	Total	Lower Division	Upper Division	Graduate
1965-66	81%	85%	54%	0%
1967-68	67%	83%	50%	8%
1969-70	66%	83%	46%	24%*

***This includes graduate students who serve as Resident Preceptors.**

TABLE 7

Percent of Average Enrollment, by Level for Selected Years

Year	Total	Lower Division	Upper Division	Graduate
1965-66	100%	81%	19%	0%
1967-68	100%	55%	42%	3%
1969-70	100%	50%	44%	6%

Of particular note is the consistently high occupancy for lower division students. Table 6 shows only a 2 percent decrease over a five-year period. Upper division students have similarly retained a rather stable proportion of oncampus residents. As graduate student enrollment has increased, the percentage of graduates living in University housing has also increased. The total proportion of students living in residence halls over the five-year period decreased from an unnatural high of 81 percent in 1965-66 to 66 percent in 1969-70.

Now notice the interdependence implicit in Table 7. The lower division students who show the highest percentage of residential occupancy have drastically decreased from 81 percent of the total enrollment in 1965-66 to 50 percent in 1969-70. In other words, the prime constituency for campus housing has decreased by 31 percent in five years. The worst is yet to come as Santa Cruz pursues its commitment to the Master Plan goal of achieving a 40 percent lower division and a 60 percent upper division ratio by 1975. The desire of junior transfers to live on campus is an unknown factor in the future, but there is a possiblility that these students could significantly raise the upper division occupancy level.

2. A decision not to live on campus is one thing, but a decision to move off campus may mean something else. Pitcher and Bosler (1970), in a survey of the 1969 graduating class, asked students to indicate why

they left the residence halls. The responses of 267 students or 71 percent of those surveyed provide the following list:

–needed to get away from campus	18%
–felt cramped by campus rules	17%
–objected to the food arrangements	16%
–couldn't stand the "dormitory" noise	16%
–wanted to save money	15%
–tired of seeing so many students	9%
–wished to join friends who were moving off campus	8%
–got married	1%

An examination of this list suggests that there are aspects of the college "communities" that impinge on the emotional life of some students and create a desire to withdraw from their pressures. The expressions "need to get away," "felt cramped," "couldn't stand the noise," and "tired of so many students" all indicate a lack of privacy, a need for solitude, and a desire to escape an environment that is too encompassing or too regimented.

Objections to the food arrangements embrace two factors–one is a constant diet of "institutionally" prepared and presented food and the other is the regimentation that contradicts an intellectual environment. Students resent the scheduled meal hours that dictate to their appetites and to their academic work and study habits.

The more "official" reasons given for leaving the residence halls are provided in two surveys of releases from residence contracts by the Office of Student Services-Student Housing (1969, 1970). The data are condensed into Table 8.

It should be noted that the respective percentages of males and females who sought release from residence contracts were:

1968-69	Men	38.5%	Women	61.5%
1969-70	Men	49.3%	Women	51.7%

The category in Table 8 which provokes the most concern is that labeled "Personal." Both the "Academic" and the "Other" category have an external or imposed quality to them. For instance, dismissal and graduation create a request for a contract release that is not entirely the choice of the student. On the other hand, the personal reasons are more student-oriented. The increase of 14.5 percent from 1968-69 to 1969-70 is examined more closely in Table 9.

TABLE 8

Requests by Percent, for Release from College Residence Contract, by Categories, for Males and Females, 1968-69 and Fall and Winter Quarters, 1969-70

	1968-1969			1969-1970		
Reasons	*Males*	*Females*	*Total*	*Males*	*Females*	*Total*
Academic[a]	34.1	36.2	35.4	28.1	34.4	31.3
Personal[b]	31.8	25.5	27.9	47.7	32.5	40.0
Other[c]	34.1	38.3	36.7	24.2	33.1	28.7
Total	100.0	100.0	100.0	100.0	100.0	100.0

[a] Includes: dismissal, graduation, intercampus transfer, exchange program, intercampus visitor, and field study.

[b] Includes: financial, live off campus, marriage, and unhappy.

[c] Includes: medical, leave of absence, withdrawal, canceled registration, and denied registration.

TABLE 9

Requests, Classified as Personal, by Percent, for Release from College Residence Contract, 1968-69 and Fall and Winter Quarters 1969-70

Reasons	1968-69	1969-70	Combined
Financial	18.7	18.6	18.6
Live Off Campus	3.1	45.2	30.9
Marriage	14.1	5.6	8.5
"Unhappy"	64.1	30.6	42.0
Total	100.0	100.0	100.0

There are several interpretations possible for the large increase in "Live Off Campus" requests. One reason might be that living on campus is becoming more unacceptable for a larger proportion of students. Another possibility is that students have found that they can get out of their contract by using this reason and do not attempt to "manufacture" a reason. Perhaps the colleges, in reviewing such requests, have found it easier to release a student from his contract than to deal with the basic reasons for the request. In any event, a category such as this which, over the combined five quarters, represents 10.8 percent of all requests granted, should be a prime source of data on student attitudes, student preferences, and student-perceived problems in residential living.

The "Unhappy" category of requests diminished by half in the two periods surveyed. It is strongly suspected that some of the increase in the "Live Off Campus" category accounts for some of the decrease here. However, for the combined five quarters, the "Unhappy" category is 14.6 percent of the total requests and demands closer inspection. Excerpts from the requests of students in this category include (Office of Student Services-Student Housing, undated):

> "need more privacy . . . greater feeling of independence . . . atmosphere is not conducive to studying . . . makes me tense, nervous . . . unable to relax enough to concentrate on my work." (male)
>
> ". . . so I can get away from being dragged down by my surroundings . . . the situation in the dorms, . . . the grossness, the noise is intolerable." (male)
>
> ". . . unhappy living the structured, routine dorm way of life . . . I just can't keep on for three more months." (female)
>
> ". . . tired of not being able to work in my own room . . . tired of trying to get enough sleep in spite of my roommate's habits." (male)
>
> ". . . I can no longer be involved in the University life 24 hours a day and be happy . . . there is just too much emotional strain on me at this point." (female)
>
> ". . . I find the closeness of quarters and lack of privacy difficult for me to live with." (female)

The greatest causes of stated unhappiness are noise, lack of privacy, food, roommate difficulties, study problems, and the emotional impact of any or all of these reasons. An additional concern is stated by members of minorities–they feel conspicuous and out of place in the residence halls.

These data reflect the unanticipated impact of approving contract releases without adequate information and without responsibility for the consequences incurred. Although the Assistant Chancellor-Student Services had countersignature, each college recommended the release of students from housing contracts without having (1) fiscal responsibility for the operation of their own residence facilities, and (2) any idea of the overall impact of releases from all five colleges. Under a decentralized plan, introduced this fall, the Business Manager-Business Services now has countersignature and each college has become responsible for the fiscal management of its residence facilities. This will not deter colleges from recommending the release of students from contracts for just causes, but it will encourage the colleges to determine with more precision the sources of discontent and to make those adjustments which are necessary for the maintenance of desired occupancy levels.

3. Data on student housing, from a slightly different point of view, were gathered by Bosler (1970) as part of a survey of 1970 UCSC graduates. Seventeen factors relating to housing were listed and graduating seniors were asked to select four factors which were important to them on each of the following:

   –Desirable oncampus features
   –Undesirable oncampus features
   –Desirable offcampus features
   –Undesirable offcampus features

   Responses were obtained from 208 students or 58 percent of the survey population. Table 10 displays the results.

   Of the oncampus features considered desirable, two items–nearness to college and nearness to students–accounted for 89 percent of the responses. This attractiveness of propinquity may merely mean convenience since only 4 percent of the respondents chose "community spirit" as a desirable feature. Four features of oncampus housing that were considered undesirable account for 79 percent of the responses–

cost, noise, food arrangements, and rules. The features of offcampus housing deemed to be most desirable are cost, be on own, need to get away, and food arrangements for a total of 82 percent of the responses. The undesirable features of offcampus housing indicated by 77 percent of the respondents were: transportation, availability of housing, nearness to college, and roommate problems. Of interest for 1970-71 will be the impact of a new city bus schedule to campus on what was the most undesirable feature of living off campus--transportation. It is possible that the unavailability of offcampus housing will become the only significant undesirable feature of living off campus. It is clearly evident as well that much needs to be done to increase the desirability

TABLE 10

Percent of Respondents Indicating the Desirability and Undesirability of Selected Housing Features

Features	On Campus		Off Campus	
	Desirable	*Undesirable*	*Desirable*	*Undesirable*
Nearness to college	66%	– %	– %	11%
Cost	2	30	38	5
Transportation	2	2	–	33
Noise	–	24	3	1
Be on own	–	3	22	–
Availability of housing	–	–	–	25
Nearness to students	23	4	1	5
Food arrangements	–	16	10	3
Need to get away	–	3	13	–
Rules	1	9	7	–
Room size	–	4	3	–
Roommate problems	–	4	–	8
Community spirit	4	–	–	1
Maids	1	–	–	2
Roommate selection	–	–	–	1
Single room	1	–	1	–
Other	–	1	2	5
Totals	100%	100%	100%	100%

of oncampus features such as cost, noise level, food arrangements, and community spirit by decreasing their present undesirable characteristics.

The problems associated with student housing are based in the facilities, the use of the facilities in the academic program, the management of the facilities (including rules and regulations), and in the expectations and desires of students. No single simplistic solution, such as required residence, will solve the problems. These problems must be considered in terms of the whole campus for only in this way will solutions arise that have been tested for their impact on other aspects of the campus. To treat the problems so that they do not create unanticipated consequences will require a meticulously accurate analysis and the clearest, most rational decision making. Both analysis and decision making are dependent upon precise information–about how things relate and about how well various aspects of the campus perform in relation to their intended purposes. Such information can only be obtained through a continuous process of monitoring all aspects of the campus that bear upon the academic, personal, and social goals for students in a residential college.

As further decentralization occurs on the campus, each college will be expected, with whatever assistance is needed or desired, to devise appropriate methods for studying their own residences. It is anticipated that students and faculty will play a major role in providing the necessary information, in devising the possible alternatives, and in evaluating the outcomes of instituted changes.

A major step toward this end has already been taken by the employment of a bursar in each college. Among other responsibilities, he will "manage" the residential houses, instead of a central housing office. Close contact with students, faculty, and college administration will produce programs and services more in keeping with the needs of the particular students in each college.

Which Model for Evaluation at Santa Cruz?

The underlying assumption thus far is that Santa Cruz should be evaluated. However, the term evaluation is being used for want of a more appropriate term. As Dressel points out in the first chapter of this monograph, evaluation has many connotations, implications, and formulations. It can be "summative" or "impressionistic"; it can be done to, for, or with; it can be concerned with costs, impacts, or patterns. At best, evaluation in the normal sense is fragmentary, judgmental, and myopic.

At its worst, evaluation in the normal sense can be destructive, lead to spurious conclusions or serve as "a form of persuasion directed at powerful people who make decisions and control resources [Trow, 1969, p. 6]."

The tasks confronting Santa Cruz are to understand what is meant by the term evaluation and to formulate a theory and methodology that best serve this particular campus.

Definitions of evaluation show little variation. Dressel (1961) states that "Evaluation involves judging the worth of an experience, idea, or process. The judgment presupposes standards or criteria [p. 6]."

Tyler (1950) indicates that the purpose of education is to produce changes in the behavior of individuals and that evaluation is "the process for determining the degree to which these changes in behavior are actually taking place [p. 69]."

Berty (1968) reiterates the notion that evaluation is the determination of the degree to which objectives are being met. He then recommends that "product evaluation"–i.e., the exhibition of performance that indicates the successful achievement of the objectives–holds the greatest potential for improving education.

This general approach to evaluation has been translated into and constrained by the scientific or basic research model. Typically, clever designs are developed that make explicit specific objectives or values, that define criterion variables, that measure how much of what is valued presently exists, that control (often only statistically) the "treated" and "untreated" individuals or programs, that produce several pages of tables (preferably at an appropriate level of statistical significance), and that are written up, published, and quoted by several generations of succeeding authors. The complications and difficulties encountered in each step of this model are overwhelming and even if moderately successful, the results are so related to the specific project that generalization is presumptive.

Social scientists have attempted to develop more appropriate models by recognizing both the evaluation aspect and the research aspect through what has been called "evaluative research." In Herzog's (1959) model for evaluating psychosocial change in individuals, the emphasis is placed on the development of criteria against which to measure or evaluate results. However, the conclusion is that evaluative research is particularly difficult to undertake, that it is costly, that existing records and data are usually inappropriate, and that it should not be undertaken if there is another form of research that could be used.

Suchman (1967) has made a clear distinction between basic research and evaluative research in terms of the purpose to be served. Basic research implies the discovery of knowledge and the success of the research is determined by the scientific validity of its finding; i.e., the finding is judged against "the rules of scientific methodology." On the other hand, evaluative research serves the purpose of learning the extent to which a program or procedure is producing some desired result. The success of evaluative research is judged by the usefulness of its finding to the administrator in improving the program or procedure. While this distinction is most useful, Suchman (1967) immediately returns to an old familiar definition of evaluation: "the determination . . . of the results . . . attained by some activity . . . designed to accomplish some valued goal or objective [p. 32]."

Along with this definition is the reliance on the standard research model but with a recognition of the long-range, intermediate, and immediate nature of different objectives. The categories of evaluation are much more related to program management and evaluation than are other models as evidenced in the use of terms: effort, performance, adequacy, efficiency, and process.

It is becoming quite clear that these research-based models are less than adequate for use at Santa Cruz or perhaps for any dynamic situation. Stufflebeam (1967) states one reason this way:

> The application of experimental design to evaluation problems conflicts with the principle that evaluation should facilitate the continual improvement of a program. Experimental design prevents, rather than promotes changes in the treatment because treatments cannot be altered in process if the data about differences between treatments are to be unequivocal [p. 128].

Further recognition of the dynamic aspects of research on innovative programs is expressed by Trow (1969) when he recommends that such research be perceived as "illuminative" rather than "evaluative" because "the value of innovation comprises the rewards gained by the faculty members who create it and are not confined to its easily measured outcomes [p. 26]."

In a response to Trow (1969), Litwak (1969) states that there is no one ideal model of evaluation and urges researchers to explore the concept of a multiple model. His tentative formulation produces 24 possible approaches to evaluation based on realistic considerations and constraints. The appropriate approach to evaluation would be determined by such factors as the complexity of the stimulus (either simple and not causally related or complex and causally related); the control that the researcher has on the stimulus in either case (complete, partial, or none); knowledge of the ends and means involved (good knowledge and good operational measures or poor knowledge and poor operational measures); and whether the design is costly or not costly. This formulation is cognizant of both the quantitative and the qualitative aspects of evaluation and encourages consideration of an appropriate approach to a wide variety of situations instead of forcing the situation to a standard, and most often, inappropriate approach.

Knowles (1970) provides insight that seems most relevant to the evaluation of innovations. He discusses the evaluation of adult learning as a "rediagnosis" of the original needs for the learning. In terms of Santa Cruz, the implication is that the extensive diagnosis of student, faculty, and institutional needs that went into its conception and founding should be "rediagnosed" in the light of present needs. The process of "rediagnosis" is less threatening than evaluation, involves the participants in the process, tends to encourage a new cycle of activities, and avoids adherence to and energies wasted on outmoded pursuits.

The insight of Knowles (1970) is rather congruent with certain aspects of system theory. An organization is an "open system," that is, its functioning is dependent upon a variety of transactions with its environment. The proper functioning and the survival of the organization is much less dependent on the *evaluation* of its various parts than it is on the accurate diagnosis of what it needs from its environment and what the environment wants or will accept from it. An organization achieves a "steady state" through the information that it collects, codes, and utilizes. When the information system of an organization is impaired or is too selective, the organization loses its ability to adjust to its environment and its steady state vanishes.

Another aspect of system theory that is relevant to this discussion is the principle of "equifinality." This principle postulates that "a system can reach the same final state from differing initial conditions and by a variety of paths

[Katz & Kahn, pp. 25, 26]."[8] This principle reinforces the importance of the need for an adequate information system, the necessity for continuous "rediagnosis," and the ability and willingness to adjust or change paths on the way to a desired final state.

The intent of this discussion has been to divert attention away from concepts of evaluation that are bound in traditional research methods and to direct attention toward a dynamic system concept. What is lost in abandoning the research model are the unrealistic notions of replicability (possible only in highly controlled laboratory experiments anyway), comparability (fallacious in terms of complex social systems), and the scientific stamp of approval (less than relevant to other than controlled scientific experiments). The scholarly bases, the intellectual rigor, and the careful design of appropriate models with checks and balances are not abandoned but are redirected to the purpose of providing the complex organization and its various components and constituents with the information necessary for making appropriate decisions and adjustments.

Guba (1968) has presented a framework for a possible model of this sort. He defines evaluation as "a process of providing and using information for making educational decisions [p. 11]." He further describes evaluation as "continuing," "multifaceted" (using many methods and techniques), "practical," and "relevant." These characteristics make evaluation a tool which provides information for judgments rather than being a judgment in itself.

Two types of evaluation are described by Guba (1968). One type he labels "homeostatic" or "congruence" evaluation. This type of evaluation is based on previously defined goals and limits and its purpose is to keep the system on target. Any deviation in the system prompts an analysis of the problem and stimulates an adjustment that is intended to put the system back on target. The second type is labeled "neomobilistic" or "contingency" evaluation. Here the purpose is to move the system out of equilibrium or to bring about change. Since this type of evaluation is not constrained by previously defined goals and limits, any deviation in the system can be analyzed, related to a multiplicity of goals or purposes, and a decision either to return the system to its former state or to permit the system to move to a new state is an open question.

---

[8] This quote is part of the excellent discussion of organizations as open systems in Daniel Katz and Robert L. Kahn, *The Social Psychology of Organizations* (New York: John Wiley and Sons, 1966).

Both types of evaluation are based on a continuous monitoring of as many facets of the system and its environment as possible. This monitoring process can provide a snapshot of the system or of its individual facets whenever necessary. Analysis of these snapshots on a regular schedule will indicate deviations and changing patterns, and will highlight aspects of the system in need of further analysis and direction through appropriate decisions. Most of the information storing, organizing, and reporting can be handled by computers.

The process may be illustrated by a cyclical pattern of activities that moves through time. There can be as many cycles as there are facets of the system to monitor and the cycles can be of differing time spans depending on the nature of the particular cycle. Figure 1 is a simplified illustration of one possible cycle that relates to the earlier analysis of students entering Santa Cruz.

The stages of the cycle require some elaboration.

Decisions on desirable admissions policy. This stage demands explicit statements of policy as well as an examination of the values, attitudes, educational philosophy, and anticipated goals that underly the policy.

Collection of information on actual admissions. This stage requires the acceptance or development of appropriate measures and procedures to assure that information relevant to the policy is obtained in a concise and exacting manner. Information unrelated to the policy may be collected when desired for specific purposes.

Organization of information into a desired format. This stage involves sorting or organizing the information into categories or combinations that summarize, average, compare, or correlate the information in a manner that has been determined by the needs and constraints of the policy and the collection process.

Report of the information. This stage merely provides a hard copy of the information organized in the previous stage. This report should accurately represent the information in its desired form and should be the basis on which all those concerned make their initial observations and comments.

Analysis of the information in the report. Although analyses of one sort or another may be undertaken by each individual who sees the report, there should be a team of specialists which critically examines the information, its

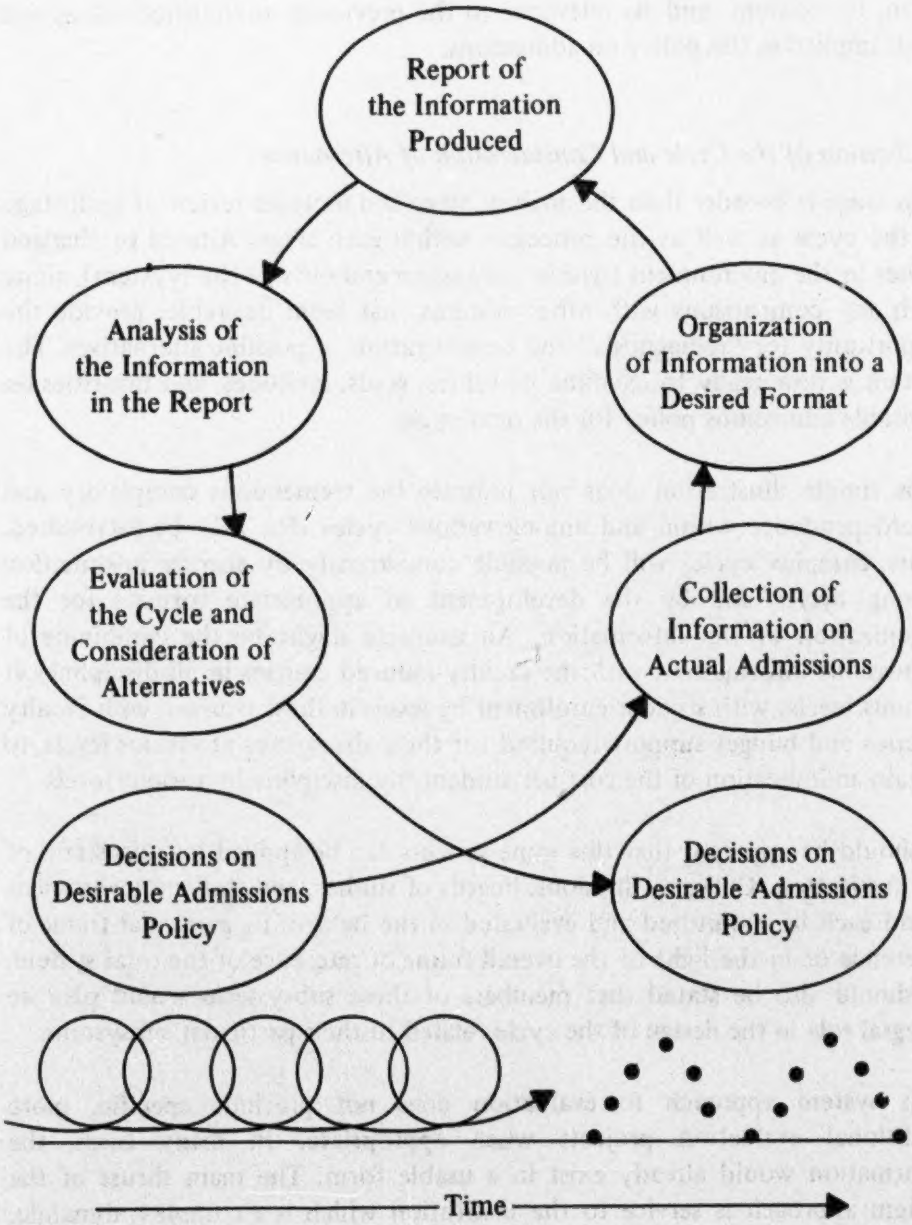

Report of the Information Produced
Analysis of the Information in the Report
Organization of Information into a Desired Format
Evaluation of the Cycle and Consideration of Alternatives
Collection of Information on Actual Admissions
Decisions on Desirable Admissions Policy
Decisions on Desirable Admissions Policy
Time

Fig. 1. Diagram to illustrate the cyclical nature of continuous evaluation, using admissions as an example. The time aspect of this process is illustrated by the series of loops at the left. However, if plotted over time, the various stages would produce the wave-like pattern at the right. In other words, the cycle does not go backward in time as suggested by the simple diagram.

form, its content, and its relevance to the previously established values and goals implied in the policy on admissions.

Evaluation of the Cycle and Consideration of Alternatives

This stage is broader than the analysis stage and includes review of each stage of the cycle as well as the processes within each stage. Altered or changed values in the environment (within the system and outside the syystem), along with any comparisons with other systems that seem desirable, provide the opportunity for "rediagnosis" and consideration of possible alternatives. The system is now ready to examine its values, goals, attitudes, and priorities on desirable admissions policy for the next cycle.

This simple illustration does not indicate the tremendous complexity and interdependence within and among various cycles that may be established. More complex cycles will be possible concurrently by sharing information among cycles and by the development of appropriate formats for the organization of the information. An example might be the combining of admissions information, with the faculty-induced courses in all disciplines at various levels, with student enrollment by levels in these courses, with faculty salaries and budget support required for these disciplines at various levels, to obtain an indication of the cost per student, by discipline by various levels.

It should be apparent that this same process can be applied to subsystems of the institution. Colleges, divisions, boards of studies, and student government could each be monitored and evaluated in the light of its particular frame of reference or in the light of the overall frame of reference of the total system. It should also be stated that members of these subsystems would play an integral role in the design of the cycle related to their particular subsystem.

This system approach to evaluation does not preclude specific, more traditional evaluation projects when appropriate. In many cases, the information would already exist in a usable form. The main thrust of the system approach is service to the institution which is a complex, dynamic, open system. The approach does not deny the desire of the institution to justify itself or to compare itself with other institutions but it does recognize that no matter how carefully controlled, results from research-based studies seldom become part of the system because they are time-bound, artificial measures of the real system. The broad participation of all levels of the system in the system's information and evaluation process is a positive inducement to recognize the need for change and to support generally agreed upon plans for change.

Development and maintenance of an evaluation system of this nature demands an organizational structure that permits an easy and trusted relationship with all aspects of the campus. Highly skilled systems/research people will be necessary in the design of various elements of the system and individuals with well-developed human relations skills will be essential to assure that the personal and interpersonal roadblocks to developing trust and honesty are understood and given perspective. The quality of the evaluation system and the quality of its relationships with various parts of the campus will largely determine the quality of the information obtained for decision making at all levels.

This discussion represents a possible approach to evaluation at Santa Cruz. Only the bare conceptual framework is presented here and considerable clarification and specification will be required before it will be offered for serious discussion. Many levels of the campus will of necessity be involved in such a process and a genuine commitment will be needed on basic principles. There must also be a major commitment of resources and staff time to the development and maintenance of such an information/evaluation system. This approach cannot be a one-shot effort. At this time, however, it appears that this approach to evaluation will best serve the needs of the campus and that it is more likely to be accepted by the campus community than other possible models.

References

American Council on Education, Office of Research. Summary of data on entering freshmen, University of California, Santa Cruz. Summary Reports for 1966, 1967, 1968, 1969.

Berty, E. Some principles and practices of evaluation. Paper presented to the West Virginia State Department of Education In-Service Program, November 1, 1968.

Bosler, R. Survey of 1970 graduates, spring 1970. Santa Cruz: University of California, 1970. Unpublished report.

Cowell College students. *Solomon's house.* Felton, Calif.: Big Trees Press, 1970.

Dressel, P. L. The essential nature of evaluation. In P. L. Dressel & Associates, *Evaluation in higher education.* Boston, Mass.: Houghton Mifflin, 1961.

Gaff, J. G., & Associates. *The cluster college.* San Francisco: Jossey-Bass, 1970.

Guba, E. G. Evaluation and change in education. Paper prepared for the Elk Grove Training and Development Center Spring Evaluation Conference, Arlington Heights, Illinois, May 16, 1968.

Heist, P., & Bilorusky, J. A special breed of student. In J. G. Gaff & Associates, *The cluster college.* San Francisco: Jossey-Bass, 1970.

Herzog, E. *Some guidelines for evaluative research.* Washington, D.C.: Children's Bureau, Social Security Administration, U.S. Department of Health, Education, and Welfare, 1959.

Katz, D., & Kahn, R. L. *The social psychology of organizations.* New York: John Wiley and Sons, 1966.

Knowles, M. S. *The modern practice of adult education: Andragogy versus pedagogy.* New York: Association Press, 1970.

Litwak, E. Comments on Professor Trow's paper entitled "Methodological Problems in the Evaluation of Innovation." From *Proceedings of the Symposium on Problems in the Evaluation of Instruction,* University of California, Los Angeles, December 1967. Los Angeles: University of California, CSE Report No. 32, May 1969.

Martin, W. B. *Conformity: Standards and change in higher education.* San Francisco: Jossey-Bass, 1969.

McHenry, D. E. University of California, Santa Cruz: Undergraduate program 1965-1966. Introductory statement in the first Santa Cruz catalog, 1965.

Office of Student Services-Student Housing. Summary of college residence contract cancellations, 1968-1969. Santa Cruz: University of California, October 9, 1969. Xeroxed report.

Office of Student Services-Student Housing. Summary of college residence contract cancellations from September 1969 to beginning of spring quarter 1970. Santa Cruz: University of California, April 13, 1970. Xeroxed report.

Office of Student Services-Student Housing. Breakdown by college of the category "Unhappy" on releases from residence contracts 1968-1969. Santa Cruz: University of California. Undated Xeroxed report.

Pitcher, A., & Bosler, R. Study of the 1969 UCSC graduates. Santa Cruz: University of California, January 1970. Mimeographed report.

Siegel, S. *Nonparametric statistics.* New York: McGraw-Hill Book Company, 1956.

Stufflebeam, D. L. The use and abuse of evaluation in Title III. *Theory into Practice,* **6**(3), June 1967, 126-133.

Suchman, E. A. *Evaluative research.* New York: Russell Sage Foundation, 1967.

Trow, M. Methodological problems in the evaluation of innovation. From *Proceedings of the Symposium of Problems in the Evaluation of Instruction,* University of California, Los Angeles, December 1967. Los Angeles: University of California, CSE Report No. 31, May 1969.

Tyler, R. W. *Basic principles of curriculum and instruction.* Chicago: University of Chicago Press, 1950.

University of California, Santa Cruz. A provisional academic plan for the Santa Cruz campus, 1965-1975. Mimeographed statement, 1962.

University of California, Santa Cruz. Cowell College: A statement of aims. Mimeographed confidential draft, November 1964.

University of California, Santa Cruz. Santa Cruz campus academic plan, 1965-1975. Mimeographed statement adopted by the Regents, January 22, 1965.

University of California, Santa Cruz. Proposed academic plan for the University of California at Santa Cruz, 1970-80. Mimeographed paper submitted June 3, 1970.

Chapter 10

JAMES MADISON COLLEGE[1]
MICHIGAN STATE UNIVERSITY

Herbert Garfinkel

Each year since the fall of 1967 James Madison College has added a freshman class and launched a new component of its projected four-year curriculum. The 1969-70 school year was Madison College's and the Class of '71's junior year. However, a milestone was reached as one member of that first entering class became the first graduate at the spring 1970 Commencement. For most of the Class of 1971, graduation will come in spring 1971. Planning for the senior year seminars, which will complete the four-year academic program, was well-advanced during this past year.

During the year just concluded, the junior year program was made operational for the first time. This included rounding out the core programs in the College's multidisciplinary fields of concentration as new junior level courses were taught. An outstanding innovation was the Madison Field Experience Program. Every member of the junior class undertook a one-term offcampus assignment in a governmental or private organization engaged in policy-making processes related to the Madison curriculum. An expanded measure of curricular flexibility was provided as students were enabled to propose an increased number of Special Topics Courses (MC 290). Also expanded was the Madison cocurricular series of programs complementary to

---

[1] This chapter is a somewhat modified form of the 1969-70 Annual Report of the College. In particular, names, recognition of individual efforts, and other matters of purely local interest have been deleted. Many persons in the College contributed drafts from which this report was prepared.

course instruction, though not for credit. And this was the year in which Case Hall was really transformed into a residential college as the balance tipped both quantitatively and qualitatively to produce a small college environment.

Student unrest pervaded all of academia in 1969-70, and our College could not but be affected in some measure. Not that our students succumbed to extremism, although many engaged in protest and participated briefly in the student strike. Few classes were affected for any length of time, and faculty members behaved in a professionally responsible manner. But the sense of unrest weakened academic values and hampered enforcement of academic standards. Nonetheless, standards were maintained, even as these were subjected to intensive questioning in an attack on all curricular requirements as inherently coercive. No doubt such questioning is necessary, when constructive in spirit, and the Madison experimental design has provided for a thorough evaluation in the fifth year of operations. Data necessary to such a systematic evaluation have been collected since the beginning. Next year, Year Four, will see a marked increase in these efforts as the College achieves its full four-year program.

The Curriculum: 1969-70

Our distinctive policy problems approach to social science teaching combines student concerns about contemporary social problems, a systematic multidisciplinary application of relevant social science disciplines, and direct student involvement with policy-making agencies as basic thrusts. This year, the major curricular development has been the successful implementation of the Field Experience Program, although a number of new courses have been taught for the first time in the College's five fields of concentration and ongoing programs have been reviewed and improved. This past year has witnessed a critical dialogue about the College's curricular focus.

General Education in James Madison College

The College has continued to provide the University's general education requirements within the broad framework utilized last year. American Thought and Language (ATL) and An Introduction to the Study of Policy Problems (the College's substitute for the University College Social Science sequence) were offered as Madison courses, while students enrolled for their natural science requirements in the University College department. Some Madison students satisfied their humanities requirement by enrolling in special Madison sections offered in University College. Special sections of Madison ATL were again offered by an excellent faculty, assisted by three writing coaches who graded papers, gave occasional lectures, and generally assisted in classroom instruction.

Continued efforts were made to experiment with the substance of the ATL sequence by the informal offering of various "tracks" emphasizing American history, humanities, and literature to take account of the differing needs and interests of Madison students. To some extent "Black" literature was emphasized in all the "tracks." Classroom instruction was also supplemented by an extensive cocurricular program, including films, panel discussions, and illustrated lectures. Overall, the Madison experiment with ATL has been successful. Continuing efforts to review the course by a joint faculty-student committee has resulted in proposals for three new "tracks" for the 1970-71 academic year, covering Radical Thought, American Humanities, and Black Literature, as well as the continuation of a related cocurricular program. A detailed review of the total program is planned for next year.

This year the College again offered special sections of the three-term humanities course sequence (Hum 241, 242, and 243) in Case Hall which experimented with the introduction of music into the humanities program. These efforts were ably assisted throughout the year by two graduate assistants who are Music Department doctoral candidates. In addition to normal duties, these assistants organized an extensive cocurricular program of listening concerts and 10 "live" performances by MSU music groups in Case Hall. These in- and out-of-class innovations were drawn on heavily as ingredients of the Humanities Department's new music "track" to be offered next year.

While the Case Hall humanities sections were warmly received by many Madison students, their enrollment figures have proved to be too small to enable the Humanities Department to continue this experiment next year. The major cause of this situation was probably the combined effects of scheduling conflicts of the humanities sections with Madison courses, the relatively heavy academic demands of the Madison sections, student concern about their specific preparation for the Humanities Department's common examination, and a lack of interest in the music focus by some students. However, next year in both the fall and spring quarters, one Case Hall section of the humanities sequence will be offered for which Madison College students will be accorded enrollment priority.

The Freshman Policy Problems Course

The unique component of the College's first-year program in general education is still the three-term sequence (An Introduction to the Study of

Policy Problems, MC 200, 201, 202) which serves both to satisfy the University's social science requirement and to introduce students to the substance of the College's five fields of concentration as the core of its upper division academic program. The course seeks to achieve a diverse set of multiple goals for James Madison freshmen: (1) to provide them with a first introduction to the major social problems (policy issues) of our day; (2) to do this from the perspective of the social sciences; (3) to give them a first nodding acquaintance with the distinctive concerns, methods, and data of the several social sciences; (4) to make a beginning in the long process of cultivating the kind of analytic reasoning skills needed for the systematic analysis both of public policy issues and of the decision-making structures and processes involved; (5) to cultivate sensitive, imaginative, and rigorous philosophical commitments and value priorities; and (6) to develop intellectual and moral orientations that will lead to coherent and productive habits of blending philosophical-ideological, scientific-objective, and concrete policy-formulation concerns.

The experience with the course during the past year was generally successful. Some of the major concerns generated by both faculty and students in their continuing evaluation have been reflected in various proposals, including: to commence the sequence with an extensive review of major philosophic issues, to widen the emphasis placed on the institutions and dynamics of the policy-making process, or to extend the coverage of specific policy issues of contemporary significance. These "grand design" issues have raised the inevitable problems of proper balance between abstract concepts and concrete issues, the coverage of a wide variety of contemporary topics at the expense of a detailed understanding of the empirical content of particular problems, as well as continuing discussions about the content of the major foci of the sequence.

In spite of these considerations, the structure and content of the freshman Policy Problems course has weathered the test of time. Moreover, the course has been generally well received by the bulk of the James Madison freshmen as a serious and well-planned introduction to social science policy problems. For example, in a College-conducted course evaluation during the fall quarter, 76 percent of the respondents indicated the course had increased their insight into social issues, with 39 percent reporting that such insight had been increased a great deal. Alternatively, 86 percent noted that the course required more preparation time than their other freshman courses, with 30 percent of all respondents reporting that they spent 10 or more hours per week on class assignments.

College Methodology Requirement

The College has continued to require at least one course in quantitative methods of the social sciences. Most Madison students take MC 205, Methodology for Policy Sciences. The course provides students with reasonable proficiency in the following areas: (1) the design and carrying out of research relevant to making enlightened social policy decisions; (2) the evaluation of research studies and proposals in terms of their ability to answer relevant questions unambiguously; and (3) the critical assessment of the validity and interpretation of data used to support arguments favoring specific policy decisions.

A major failing in the College's academic program has been the limited extent to which, in practice, the methods requirement has been woven into the fabric of the College's general curriculum. A study of the records of the College's current junior class indicates that almost one-third of these students have not yet taken the methods requirement, while approximately another one-third deferred this coursework until their junior year. Issues of major curricular concern for next year are to consider the need for an expansion of the methods requirement, the necessity for students to take this coursework early in their program, as well as the use made of their methodological competency in the total curricular program. The content of the program, grading systems, and opportunities for additional methods coursework beyond the requirement are also likely to be important issues. However, the College has already provided for additional coursework in methods as an option in its new cognate option substitute for the foreign language requirement.

Foreign Language Requirement

The College's foreign language requirement underwent a detailed review and revision during the 1969-70 academic year because of concerns by both faculty and students about the educational relevance of the College's current requirement. In part, these concerns related to a questioning of the effectiveness of only one year of foreign language study, plus related area courses, the option generally chosen by the bulk of students, as well as the desirability of providing an option for students to study social science methodology rather than a foreign language. Some students objected to the language requirement as a part of their general rejection of all academic requirements, arguing that students should be given the maximum flexibility in developing their own academic programs.

In response to these criticisms, the College Curriculum Committee appointed a faculty-student subcommittee which presented a proposal for a substitute liberal arts cognate which was approved by the College Curriculum Committee.

Cognate Option: Foreign Languages, Area Studies, Humanities, or Social Science Methods

All students will be required to complete *one* of the following options:

A. Attainment of second-year competency in a foreign language. This may be met by satisfactory completion of University courses or satisfactory performance on a placement examination. No credits toward graduation will be granted for completion via the placement examination.

B. Satisfactory completion of six courses in foreign area study as follows:

   1. Six courses in foreign area study, or

   2. Three courses in each of two foreign areas.

   3. Only courses listed at three credits or above are acceptable.

   4. Courses should be selected by students in consultation with their academic adviser to provide an integrated, systematic program of study. In addition, it is strongly recommended that students consult with their advisers to select academic programs which maintain a proper balance between lower level and upper division courses.

C. Satisfactory completion of six courses in humanities as follows:

   1. Courses selected by students must have the written approval of their advisers. In addition, it is strongly recommended that students consult with their advisers to select academic programs which maintain a proper balance between lower and upper division courses.

   2. Only courses listed at three credits or above are acceptable.

D. Satisfactory completion of six courses in methods applicable to the study of social science. Students may choose from courses listed in the various units of the College of Social Science and in the Departments of Computer Science, Economics, Mathematics, Philosophy, Statistics, and Systems Science.

This proposal, now having been approved by the Madison faculty, must be referred for final approval by the University Curriculum Committee and the Academic Council.

The Annual Writing Requirement

The College has continued its Annual Writing Requirement under which all of its students must complete at least one substantial documented paper annually. Students may satisfy this requirement with papers written as an assignment for a Madison course, or a general University course, or by written work satisfying the requirements of an independent study course. In addition, the James Madison College writing coaches are available to assist students as they prepare papers to satisfy this requirement. Failure to satisfy the requirement, without official approval of a deadline extension, results initially in students being put on academic probation and ultimately in their withdrawal from the College.

During the year, the Annual Writing Requirement has been protested by some Madison students for various reasons: some students used this issue to register their opposition to all academic requirements, while a probably larger group argued that either the writing assignments normally required of them in their courses made the general requirement redundant or that alternatively, their assignments did not provide opportunities to satisfy the writing requirement without extensive expenditures of additional time and effort. A number of students have also suggested that the criteria applied to the requirement by some faculty members have been so lax that efforts to implement it seemed irrelevant. Sporadic student efforts have been made during the year to organize a boycott of the writing requirement. Beyond this, some of the College's faculty have raised questions relative to the procedures designed to implement the writing requirement and, in fact, a significant percentage of Madison College courses have required writing assignments which do not readily enable students to satisfy the writing requirement.

In response to these general considerations and criticisms, the faculty approved the following resolution during the spring term:

> **Freshmen students may satisfy the Writing Requirement by several short papers as well as by the present 10-page paper. It is understood that these papers will be certified by faculty members as representing writing of the required level of excellence and the student's real ability.**

However, this resolution applies only to the current academic year, and the principles underlying the Writing Requirement as well as the means used for its implementation for all Madison students will be subjected to a detailed review in the College during the next academic year. Accordingly, until such time as the current requirement is reaffirmed, modified, or abolished by the faculty, it, together with the sanctions stipulated for noncompliance, will be maintained.

Field of Concentration Core Programs

In the sophomore year each student chooses to pursue one of the College's five field-of-concentration core programs. Each field is comprised of courses drawn from the University at large as well as Madison College and in accordance with the College's general guideline that approximately 50 percent of the total 180 credit hours required for graduation be taken in the College, an appropriate balance of Madison College to University-wide courses is provided in each field. In addition, the College faculty recently approved an option enabling students to take up to 75 percent of their total credits in Madison courses if they choose to do so. Students may use free electives beyond the core program requirements to earn a teaching certificate, pursue coordinate majors, or merely satisfy general educational interests in the social sciences or any other academic area.

The fields of concentration focus on specific policy problem areas, but are multidisciplinary in approach; they lean heavily on MSU courses in relevant disciplines. Each field of concentration is to build on courses already available in the social sciences. Thus, the core of each program usually consists of a sophomore-level introduction to the contemporary policy issues of the particular field, a middle set of courses which draw upon the contributions of related disciplines, and a senior year reexamination of selected policy issues in light of the students' field experience.

During 1969-70 the following five fields were again offered in the College: (1) Ethnic and Religious Intergroup Relations Policy Problems, (2) International Relations Policy Problems, (3) Justice, Morality, and Constitutional Democracy Policy Problems, (4) Socioeconomic Policy Problems, and (5) Urban Community Policy Problems.

The experience with the academic programs in the fields has been mixed, although all new courses were implemented as planned during the year. Some problems were a result of teaching new courses for the first time, as well as of the difficulties associated with effectively coordinating the focus of specific courses as ingredients of a particular field's general program. Beyond these considerations, a marked enrollment imbalance has emerged in the academic program choices of James Madison College students. The JMCD and Ethnic and Intergroup Relations Fields of Concentration have attracted 70 percent of the College's total junior class enrollment. A similar pattern, although with less extreme enrollment variations, has emerged for the current sophomore class. A variety of factors appears to be responsible for this enrollment pattern, including variations in the quality and difficulty of the program offered in the fields, the more general attractiveness to many Madison students of a humanistic-philosophical curriculum rather than a "harder" social scientific approach to the study of policy problems, and a general continuing interest in and concern about the problems of American race relations. The quality of instruction offered in the various programs has also influenced these enrollment figures.

The specific division of the College's general curriculum into five fields of concentration also has been questioned. Some students, seeing the close linkage between the problems of race relations and those of the central city, have urged an amalgamation of the Urban and Ethnic Fields. This is resisted by those concerned to retain the emphasis on religious and national minority group problems along with that concerning race relations and Black Studies. Thus, while the third year of the College's program has been implemented with relative success, significant questions relative to its general focus and orientation have been raised and must be confronted during the 1970-71 academic session. Evaluations of each of the five areas have been made and are here summarized.

Ethnic and Religious Intergroup Relations Policy Problems

Generally speaking, the original curriculum has proved sound and innovative, but there is a need to provide a closer coordination between the various courses in the program. Infrequent meetings of the faculty-student field committee and the lack of cocurricular programs related to the curriculum have reduced the effectiveness of the program. Urgent future needs are the recruitment of both a senior qualified faculty member as field chairman and more Black faculty. In addition, greater emphasis needs to be placed on a wider comparative focus, both domestically and internationally, for the problems of ethnic and religious intergroup relations.

International Relations Policy Problems

Major problems include the need to revise the curriculum to better integrate its focus on world politics within a context of detailed considerations of the structure and process of American foreign policy together with greater stress on comparative economic and political systems and to recruit to the program a permanently committed staff with more diverse fields of interest beyond the problems of development and East Asia.

Justice, Morality, and Constitutional Democracy Policy Problems

Major issues emerging in a review of the field are the need to achieve a better coherence in the curriculum, greater flexibility and more innovation in teaching methods, and more emphasis placed on the problems of modern social theory and the governing of a constitutional democracy.

Socioeconomic Policy Problems

There has been general agreement that the goals and objectives of the field are still appropriate to the College's program. In addition, the separate curricular tracks available to students with interests in economic and social welfare policy have been continued. However, a major concern has been the inability of the field to attract a larger student enrollment. Among the many factors likely to have discouraged enrollment were: the course content and a lack of coordination in various aspects of the program, the absence of Madison courses in the field at the sophomore level, its larger number of required credits in comparison with other Madison programs, the limited specificity of the program to students largely interested in a social work curriculum, and the heavy emphasis on the study of economics as a major part of the curriculum.

A review of the field, with particular emphasis on the economic "track," is planned for next year. The field committee has met frequently and has acted as an important deliberating body. However, a high rate of student turnover in the committee's membership has reduced effective participation in these discussions.

Urban Community Policy Problems

Major problems in the field were the result of the lack of a balanced junior-senior staff committed to urban studies, limited coherence among the disciplines represented in the program, and the general absence of a student body with a career or graduate school orientation to urban studies. The field committee met infrequently, with the students often taking the initiative in identifying their own concerns, but in the absence of an operational forum to respond to these problems, their efforts were often ineffectual.

The Field Experience Program

As part of their academic program, all Madison students are required to enroll for one quarter (15 academic credits) of field experience, usually during the spring term of their junior year. However, a longer program or a different scheduling sequence may be approved in appropriate circumstances. The program is an integral part of the College's curriculum; students are placed in agencies or institutions to bring them into direct contact with practical situations relevant to questions of public policy. Participating students are graded on a pass-no grade basis and are required to complete an extensive journal including both critical analyses and evaluation as well as the recording of their general experiences in diary fashion. An important by-product of this experience is to acquaint students with career opportunities in their fields of study and participating agencies with possible candidates for full-time positions in their programs.

Starting with virtually no placements on hand, 319 positions in various agencies and institutions have been negotiated, and through the summer term, 1970, 100 students have been placed in openings in Washington, D.C., New York, Chicago, and Michigan, as well as overseas. In addition, administrative procedures have been developed for the program, and students have been oriented and prepared for their placements in the field.

Although it is too early to make any final judgments about the experience of the program, students' responses typically have been as follows: (1) Initially there appeared to be great relief and considerable satisfaction that the program was underway and offered concrete and funded opportunities as it does; (2) early in the spring the realities of the "real world" and the occasional trivia and uncertainty associated with short-term appointments have produced some criticisms of the program. Also, students have not easily made the move from their experience in the classroom to the considerably more routine and structured life of administrative officialdom. More recently, reports indicating more meaningful experiences and general satisfaction with the program have been received from many students. Most faculty response concerning the development and administration of the program has been favorable. However, while many faculty academic advisers have signed placement recommendations automatically, others have been more critical of what they believe to be a system which does not provide the adviser with enough opportunities to accept or reject a recommendation of the FEP office.

Major problems emerging in the program during this year have been those associated with the successful matching of student demand for placements with available openings or last-minute changes in plans by participating students or organizations; the unevenness of financial support for the various available positions and the lack of detailed job descriptions for particular openings; and variations in the quality and reliability of some students, resulting in part from their inadequate preparation for field experience. All these problems, together with the recurrent need to provide enough field experience openings for all applicants, are challenges for the program's administration to solve next year. Nevertheless, on every important count this major component in the College's four-year developmental program is an unqualified success.

Independent Study

The College is authorized under its MC 295 and 495 numbers to offer credit for independent study under the guidance of Madison College faculty members. During 1969-70, 83 students took MC 295 and 14 took MC 495, generating a total of 212 credit hours. Projects ranged widely and in some instances independent study provided a basis for preparing a paper satisfying the annual writing requirement. In some cases, the instructors supervising the several independent study courses brought their students together several times during the term to discuss their work on individual projects. Five students also worked together in an independent study course involving a research project in the College. Normally, however, students work on a tutorial basis with individual instructors to enable them to explore a particular topic in more depth. In addition, to be eligible for independent study, students must have a cumulative 2.0 average at the beginning of the term and be in good academic standing. Exceptions to this rule may be made only with the positive recommendation of the instructor and academic adviser and with concurrence of the assistant dean.

Coordinate Majors and Preprofessional Programs

From the beginning the College has sought to provide students with opportunities to undertake specialized courses of study in addition to their multidisciplinary field of concentration. In many instances, coordinate majors–complementary majors in the social sciences and related disciplines–are available to students who wish to obtain a more specialized undergraduate education as a preparation for graduate study. Most graduate schools do not normally require students to major in a discipline to be accepted for study if students have had an adequate alternative undergraduate education, as is provided in the Madison field of concentration core programs. Nevertheless, this year the College again has tried to work out arrangements with various departments and schools to enable students to satisfy jointly both Madison College and departmental major requirements while allowing the maximum opportunity for a well-balanced undergraduate program.

To this end, proposals for coordinate majors have been presented to the School of Social Work and the Department of Psychology and during the next academic year coordinate major proposals will be reviewed with the other departments in the College of Social Science. An agreement has already been worked out with the School of Social Work under which a limited number of Madison students will be able to jointly enroll in the School's undergraduate program. In addition, arrangements are being worked out with the Registrar providing that the transcripts of those James Madison students pursuing coordinate or dual majors will officially designate that they have completed the necessary requirements for the relevant departmental major. Finally, the College has tried to clarify the terminology commonly used to describe those programs undertaken by Madison students in addition to their multidisciplinary core programs: (1) "Cognate Programs" refers to a program of study in a particular field worked out by a student in consultation with his academic adviser; (2) "Coordinate Majors" refers to situations in which students follow to the letter the particular departmental requirements of a specific academic program as spelled out in the *MSU Catalog* and the College makes a certification to the effect to the Registrar; and (3) "Dual Majors" are programs worked out between the College and particular departments formally enabling students to pursue academic programs offered in both units.

Similar efforts have been launched with respect to preprofessional programs, although the requirements for some of them, e.g., a prelaw program, have been worked out unilaterally by the College. In particular, the College's multidisciplinary field of concentration core programs have been approved by the Teacher Education Council as appropriate courses of study leading to certification as Secondary School Social Science Teachers. Under this program students jointly enroll in James Madison and in the College of Education, and they complete various course requirements in general education, social science, history, and professional education. The experience with this program during the year has demonstrated real difficulties for students in finding enough time in a four-year program to complete both the secondary education certification and the College requirements. Accordingly, the College has made arrangements to enable students enrolled in this program to have a wider choice of free electives by allowing them to double count credits used for their required history minor and the College's general foreign areas studies and humanities cognate requirements. Currently 16 members of the junior class are enrolled in the Secondary Education Certification Program.

In conclusion, the curriculum continues to fulfill the objectives for which it was designed. Many Madison students find their interests sustained and nourished by it. This year, however, some have questioned the basic philosophy and goals of the College as well as the implementation of specific ingredients of its academic program. And these remain important challenges to be considered during the fourth and rounding out senior year of the College's developmental program.

The Residential College as an Environment for Learning

By a residential college we mean that teaching and curriculum, scholarly and cultural values are critical elements of the college where students live and study. This is more than just housing students in a building, or in a little group of buildings and trying to make life comfortable for them, or trying to develop a sense of community among them. A residential college is a place where the intellectual offerings of the college are integrated with the social and individual lives of the students. By bringing the curriculum more directly into the place where people live, we have a better chance of showing students the relevance of academic learning to their lives and of generating the kind of interest and motivation necessary to sustain them when they undertake difficult scholarly work. Understandably, people in the psychoanalytic, student-centered, or student personnel services stress the emotional and personality development of students, but there is no way to get around the fact that in a college the curriculum and teaching are the *raison d'etre* and the basic elements around which individual personalities and communal relationships develop.

One thing we are after in James Madison College is to have the benefits of smallness within a large university. If students are to gain the most from their teachers, they must see them in situations in which the teachers reveal themselves as people who can inspire by their personal dedication to scholarship and their professional commitment to students as the apprentice-participants in the collegiate living-learning situation.

Students ought to be in a situation where they can also know other students well; where they can find people who share common purposes. In the modern university, too many students go through four years of college without making a friend with whom they maintain touch for more than a month after graduating. It is difficult to establish a basis for lasting friendships, to overcome the fear that other students might take advantage of them if they reveal too much about themselves. This may be the reason students talk so frequently about different housing arrangements. They have the feeling that somehow, somewhere, if they could just shake up the housing situation enough, their sense of intimacy with other people could be improved.

A major need is to try to create a more harmonious relationship among the different parts of university experience. Too often universities separate learning from life, science from practice, this course from all other courses, feeling from thinking, youth from adults–in the interest of some good purposes, but at the expense of individual development based upon intellectual growth.

Our aim is to try somehow to restore the wholeness of this enterprise.

The Office of Student Relations

The Madison College Office of Student Relations is concerned with the goals of the Residential College which center on the personal development and academic matriculation of our individual students. The issues handled by this office range from questions of personal counseling, study habits, roommate conflicts, and career choice to more traumatic issues of identity crisis, sexuality, drugs, and other dilemmas besetting typical college students. A great deal of staff time is spent in individual and group counseling sessions.

In the fall term, 1969, the College experimented by having the dually-appointed Head Resident Advisers-Assistant Directors of Student Relations live out of the hall. The other members of the advisory staff, who are dually appointed with the Dean of Students' Office, consisted of two resident-assistant head advisers and two graduate advisers. The experiment proved unsuccessful during fall term because it was difficult to provide the day-to-day coverage of the hall activities from a live-out position. Therefore, a major reorganization took place effective winter term, 1970. As a result of two resignations within the advisory staff, it became possible to provide a full complement of six resident advisory staff people.

The joint appointments between the College and the University Dean of Students' Office are seen as a critical link between the College program and the structure in which the students live. This linkage between the academic and personnel programs of the University distinguishes James Madison College from most other academic units on this campus. Rounding out the Resident Advisory staff, the College has 21 undergraduate resident assistants (RA) serving within the building. There is one RA assigned to each "house" in the College to assist students in academic and personal matters. Their fine work at the grass roots level goes a long way toward easing the adjustment to college and the personal difficulties that beset people between the ages of 18 to 22.

This past year, 1969-70, the Office of Student Relations took primary responsibility for a drug education program. Also undertaken was the formation of a Madision Environmental Studies Interest Group (MESIG). The group met throughout the winter and spring terms discussing the problems of research and evaluation of Madison students and environment. They also spent time planning for research and evaluation programs which will take place during the 1970-71 academic year. Examples of these research proposals include personality testing, impact of environment, value orientations, etc.

Cocurricular Program 1969-70

The Cocurricular Program at James Madison College was established to supplement the normal coursework of the student and as a cultural enrichment of the residential-learning environment. The primary objective of the program was to infuse the student peer group culture with academic and cultural values.

Faculty used the cocurricular format to experiment with various teaching aids and to supplement their course material. Students made use of the College facilities through this program to engage in many of their own interests such as music programs and as an expansion of their knowledge of the contemporary world and local events. Over the year the co-curricular series, topics, speakers, films, panel discussions, and the like presented a wide array of subjects, attempting to provide all students in the College with interesting and informative programs.

Inputs into the Cocurricular Program came from several segments of the College. The Fields of Concentration in the College and the Academic and Scholastic Committee of Case Hall sponsored Cocurricular Programs. The Division of Student Relations sponsored programs on drugs, and the

Introductory Policy Problems course again sponsored a number of fine cocurricular events.

Many of the programs were organized by a Madison student-run Cocurricular Committee. The Committee functioned as a clearinghouse for ideas and requests for cocurricular events. They strove to supplement academic life by sponsoring complementary coffee hour discussions, lectures, and slide talks.

Many outstanding cocurricular programs were presented in relation to specific courses. The humanities series was enriched by 11 appearances of MSU musicians, who performed in the Madison Library and in the North and South lounges. These programs were well attended by students and faculty in the Hall. They provided a unique type of cultural enrichment which will no doubt be adopted by other residence halls.

The cocurricular program provided an important supplement this year to the Madison curriculum. It attempted to meet a broad spectrum of interests and to complement the formal curriculum. The success of the program is the product of joint planning by faculty and students, as well as flexibility to create programs on short notice when it is discovered that an interesting speaker will be visiting the campus.

To facilitate new students' orientation toward the College and to help create better faculty-student relationships, Madison students and faculty again participated in a Weekend-Away on October 10, 11, and 12, 1969. One hundred and eighty-six freshmen, sophomores, and juniors attended the weekend, as well as 17 faculty members and a few of their families. Most students enjoyed the weekend and found it worthwhile. The program included a number of small-group discussions on matters relating to college life and ample time for social activities, as well as two films. Some students and faculty members complained of indecorous behavior. The entire experience will be evaluated both as to costs and as to the defects of the program to determine whether future Weekends will be organized.

Many students expressed an interest in participating in small-group experiences run by professionally trained leaders. Ten developmental groups were initiated and run during the winter term, 1970. These groups were led by counselors from the Wonders Counseling Center and the Assistant Director of Student Relations. There continues to be high interest in this experience in the Madison environment and no doubt such developmental groups will also be created for the next academic year. Approximately 100 students participated in these groups. A systematic evaluation of this program will also be undertaken next year.

An experiment was attempted in Madison during this past academic year wherein Madison College undergraduate Resident Assistants were made academic advisers for the freshmen living on their floor. Four RAs were selected to participate and met regularly throughout the year with the Associate Dean of the College to discuss problems and answer questions. An evaluation of the experiment revealed that the students felt that having their RA as an academic adviser facilitated their personal and academic adjustment to college, although the survey showed that discussions were only infrequently directly on academic matters. However, because the RAs are so readily available on the floor, the students sought academic advising from their resident assistant more frequently than did students who had faculty advisers. This experimental program will be continued for another academic year.

The College continues to be concerned about inproving the living environment for students in the College. Carpeting would greatly reduce the noise level in the rooms and in the hallways. More flexibility is needed in the arrangements of rooms. Apartment arrangements, such as have been constructed in Fee Hall, would be important to develop in this College to encourage students to remain in the building. More single rooms should also be made available to students. In a residential college where the community is small and students are well-known to each other, there is a greater premium on privacy than in other residence halls. Students need more space in the building where they can be alone. Hopefully, more arrangements will be made to meet this need in the future.

Students

Madison students continue to be a highly interesting group of individuals. Their concern for social issues and their intellectual curiosity make them a gratifying group of students for faculty to work with. They represent a broad range of political opinion, from the very conservative to the very radical. Increasingly, however, the political bent of students in this College leans in a "liberal" direction. Some conservative students have indicated feelings of isolation at times in the College, so a concerted effort will continue to be made to seek out conservative spokesmen to visit the College and represent that political viewpoint. Our students are outgoing, verbal, academically aggressive, and questioning. They are sometimes impatient with "the system," but maintain a commitment to try to work within it.

Much more data should be gathered on the Madison student population. At the present time, we have their entrance scores as well as other general University testing results. But information is needed on their personality, orientations, values, and perceptions, as well as academic achievement. Some research was done this year by the Assistant Director of Student Relations. He administered the College and University Environment Scale (CUES) to Madison students and faculty. From this research it is evident that Madison students: see professors as dedicated scholars; value careful reasoning and logic; see their professors as helpful; are actively concerned about national and international affairs; are interested in controversial speakers; like to argue; perceive that course materials are frequently revised; share problems with each other; feel free to express strong convictions; have colorful parties; like the counseling services available in the college; and see the classroom atmosphere as friendly.

CUES data also reveal extensive faculty-student relationships in this College and that the faculty strongly support scholarly values. Generally, students perceive the environment as being intellectually stimulating and gratifying.

The Madison faculty perceptions of the environment, based on the CUES study, indicate that they sense a strong community, see a large amount of pressure for scholarship and are satisfied with the quality of student-faculty relationships. They sense a high level of College morale in this environment.

The percentage of withdrawals from the total population of James Madison College has declined during the 1969-70 year. As of July 1, 1970, 24 Madison students had withdrawn from Michigan State University, or 4.4 percent of initial enrollment, while 104 students transferred to other units in the University, representing 19.2 percent of initial enrollment. Overall, the 1969-70 transfer rate was approximately the same as the 1968-69 figure, and below that recorded for 1967-68.

Each Madison student who withdraws from the College is seen by either the Director or Assistant Director of Student Relations for a withdrawal interview. At this time the reason or reasons for the change are sought, with the hope of insuring that the student is making the best decision based on his educational needs. The two main reasons offered for leaving are a lack of flexibility in the curricular structure and a feeling that the curricular and personal bent of most of the people in Madison is too liberal, thereby not satisfying those with more conservative viewpoints. The students leaving represent a wide spectrum of academic and personal interests. As in the past students commonly transfer to other units as freshmen or sophomores because of a change in academic and vocational plans.

Madison has not lost large numbers of academically exceptional students, for only 26 percent of the total number of transfers had grade point averages of 3.1 or above. A review of the academic unit destinations of the Madison transfers shows no consistent pattern of strong or weak students choosing particular academic programs.

During the past year, James Madison College had, for the first time, a significant offcampus population. Upon attaining junior status, many students felt the need to gain greater independence and privacy by moving into offcampus apartments. In order to get a better understanding of their reasons, the Office of Student Relations conducted an offcampus survey administered during winter term, 1970. In summary, the report indicates that physical changes within Case Hall would not appear to make a significant difference, nor would a liberalization of University or management rules and regulations. It simply appears that some students do not care for institutional living. It supports the notion that students who move off campus do so because they find a residence hall atmosphere stifling and lacking in privacy. The offcampus breakdown still represents a small percentage of the total Madison population. Not including the commuters who are defined as those people living at home or who are married, the offcampus population was only 15 percent of the total enrollment in the fall of 1969.

As a result of our self-selective admissions policy, the College includes Alumni Distinguished Scholars, National Merit Scholars, as well as those students who are barely admitted with mediocre high school records. However, a review of the median test scores of the freshmen admitted to Madison College in the fall of 1969 reveals that they score higher than the average freshmen entering MSU. For example, in overall reading ability, our students scored third in a rank ordering of the colleges comprising MSU. On the SAT exams, Madison ranked fifth in total performance behind Briggs, Natural Science, Morrill, and Engineering students. On the SAT numerical section, James Madison College ranked sixth behind Lyman Briggs, Engineering, Natural Science, Agriculture, and Veterinary Medicine. As in the past, it is apparent that James Madison continues to attract a disproportionate number of high ability students.

Working with Madison students is a stimulating experience both in and out of the classroom. Generally speaking, our students are bright, articulate, aggressive, and concerned with their personal and intellectual development. Within the community, there is a great deal of sharing and mutual concern among students, faculty, and staff. While disagreements have emerged between students and faculty and administrators as to the nature of the Madison experiment, all groups share the feeling that Madison College provides uniquely meaningful academic and personal opportunities at Michigan State University.

Faculty and Staff

During 1969-70, the College's recruitment efforts were less successful than the 1968-69 academic year, although several highly qualified faculty members were added.

The College has continued to seek joint appointments with other units in the University, but during the past year these efforts have been unsuccessful, possibly because of the different emphasis of the College, stressing quality undergraduate instruction, and that of the departments which lay greater stress on the importance of research and graduate instruction. None of the College's newly appointed faculty will be joint appointments next year; and, as a result of various staff changes, only 7 of the College's 25 faculty members will be jointly appointed with various MSU departments. However, the College again intends to make every effort to negotiate joint appointments in the coming academic year.

In spite of this drift toward an increasingly small number of joint appointments, the College's faculty remains highly regarded in the general University as a group of competent teachers and scholars. A measure of this is indicated by the fact that four members of the Madison core faculty have been asked to teach in other units of the University during the coming year. However, another general recruitment problem has been the College's inability to attract social scientists with strong behavioral science orientations, probably because of the desire of scholars with such interests to be appointed in units which place heavier emphasis on research and graduate-level instruction.

Evaluation in James Madison College

As a carefully designed experimental program, evaluation is necessarily an important aspect of the Madison College experience. During the year, many of the evaluation activities related to an ongoing internal review of the College's programs; others were associated with plans for a detailed review during the College's fifth year.

With regard to its own internal concerns, the College has expanded its collection of data relative to its student body (see **Students** section of this report) and has taken the first steps toward beginning a systematic review of the impact of the Madison environment on the learning and personal development of the College's student body. Beyond this, the College has again carried out a review of its academic programs initially through the use of its own course evaluation in the fall term and subsequently through a general administration of the University's Student Instructional Rating Reports (SIRR). In response to the University's new policy on Instructional Rating Reports in December 1969, the College's faculty approved arrangements under which the College Advisory Council would be given access to these survey results as an input into their advice to the Dean relative to merit increases, promotion, retention, and tenure. The SIRR was administered in all Madison courses, and during the winter and spring terms the results provided an important input into discussions with faculty concerning their effectiveness as undergraduate instructors.

The College has also been concerned to examine the matriculation patterns of its first graduates, the class of 1971, both to ensure that they are making satisfactory progress toward their degrees as well as to evaluate the character of their academic programs. In this regard, the following are the major conclusions drawn from a detailed study of the Dean's matriculation records of the 134 James Madison juniors undertaken in the spring term of 1970.

1. None of Madison College's juniors has substantial or "unrepairable" academic deficiencies which cannot be alleviated during the coming academic year. Moreover, the overwhelming bulk of students are in "good standing" and are making satisfactory progress toward their degrees.

2. Only 60 of the 134 juniors completed the College's Methodology requirement in their freshman and sophomore years. Moreover, at the present time, 30 percent of the junior class (40 of 134 students) have not yet taken MC 205, Methodology for Policy Analysis, or an approved substitute. These data show the extent to which the methodology requirement is not being seen as an integral and cumulative part of the College's academic program and cannot be drawn on in upper division coursework in a student's Field of Concentration.

3. The bulk of the junior class has completed the language requirement. Nevertheless, 33 of 134 students are still in the process of completing it, while 16 students have not even begun this requirement as of the spring quarter of their junior year. In addition, there is fairly clear evidence that students are choosing "related area

courses" which effectively complement their one-year language study. Beyond this, however, only a tiny minority of students are taking two years of a language to satisfy the requirement, with this small total falling away almost to nothing if International Relations students are excluded.

4. Generally speaking, the general education sequences are begun and completed early in a student's academic career. However, a goodly number of students have not yet completed one or both of Humanities 243 and Natural Science 193. In addition to this apparent unwillingness to complete these general education sequences, a number of students are also deferring the completion of their physical education requirement.

5. Slightly more than half of the junior class has taken more credits in general University courses than in courses offered in the College. However, the largest single group of students–61 percent in all–has taken between 40 and 50 percent of all their coursework in the general University. A total of 25 percent, or 35 students, has taken more than 60 percent of their course credits in the general University.

6. As a general rule, students are taking all of the core requirements in their fields in a fairly systematic way. However, in some programs of study there is evidence that some sophomore-level requirements are being bypassed and are not taken prior to enrollment in junior-level courses.

7. The study shows three general patterns emerging relative to electives in the academic programs of the junior class. First, students in those fields (IR, JMCD, and Ethnic) who are required to take work in three related areas normally satisfy this requirement minimally with a heavy concentration of work in one discipline, usually in sociology or psychology. Secondly, the major source of most general electives was from the Humanities (usually in departments in the College of Arts and Letters, such as History, Art, and Philosophy) and the second choice within this elective group were courses chosen from departments in the College of Social Science. Third, apart from a limited enrollment in specific Madison electives (Education and Poverty, Contemporary Ideologies, and Utopia and the Quest for Social Progress) and Special Topics courses, Madison juniors with very few exceptions did not take Madison courses outside their own fields of study as electives.

8. Currently 52 percent of the junior class (70 of 134 students) have or will complete their Field Experience by the end of spring quarter, 1970. Roughly 30 students will enroll for Field Experience this summer, meaning that approximately 10 to 20 students will be taking Field Experience in their senior year. (Precise figures are unavailable because final summer enrollment for field experience is not yet known and some of the 134 juniors are actually "sophomores" with advanced standing; i.e., students with 85 credits or more are classified as juniors at MSU.)

Activities relative to the external evaluation of the College's program commenced during the 1969-70 academic year, although a major review of the College is not planned until its fifth year, the 1971-72 academic session. An Educational Policies Committee subcommittee on Residential Colleges and Living-Learning Units began a review at the request of the Provost of the place of these programs within the structure of undergraduate education at

Michigan State University. The subcommittee met with the Deans and interested faculty members of James Madison College on at least two occasions. At the subcommittee's request, the College developed a statement relative to a proper basis for its evaluation in which it stressed that the evaluation be based on the original mission of the College, that the review should occur in the fifth year of the program to provide experience with two graduating classes as well as opportunities for a better comparison of costs; that the College's academic accomplishments be viewed both objectively and subjectively; and that an outside board of distinguished educators be invited to participate in the evaluation. In addition, the College also prepared at the subcommittee's request a list of educational "spin-offs" identifiable in the College's program which might have applicability in the University at large.

The College has also continued to consult and cooperate with MSU's Office of Institutional Research (OIR) on its ongoing evaluation program. In particular, in the fall term the College's 1969-70 freshman class participated in a questionnaire survey developed jointly with the Office of Institutional Research which sought information on their views of University life and career goals. However, the major thrust of the OIR's efforts this year has been to develop an appropriate set of costing categories to provide a truer comparison of the costs of the residential colleges with those of other units in the University. These costing categories, unlike those generally applied in the University, specifically take account of the nonrecurrable research and development expenditures incurred in the development of the program as well as the costs of the College's Office of Student Relations; thus it has been recognized that both of these cost components are of less significance in or not applicable to other units in the University. The College has collaborated with OIR in efforts to collect the relevant data to be applied to these costing categories. Additional emphasis on such cost evaluations is planned for the coming year.

Beyond participating in this cost study, the College also has begun its own review of the overall allocation of financial resources in its program. A study of the various College program costs and costs per credit hour was undertaken, together with some preliminary efforts at a general cost-benefit analysis for the total program. More specifically, for the coming academic year, cost reductions in the College's programs are foreseen as a result of a decrease in faculty force, cost reductions in the Office of Student Relations, and substantial cutbacks in the number of graduate assistants in the College. The need to carefully evaluate the internal allocation of the College's budget and to achieve additional cost reductions remain important undertakings in the coming year. In this regard, issues such as the costs and staffing ratios in the Freshman Policy Problems course, access of non-Madison students to Madison courses, and direct College participation in expanded student recruitment will be among the objects of a careful cost evaluation.

Beyond this need to provide an effective factual basis for a review and evaluation of the College's program, some concerns also have emerged relative to the character of the College's general objectives, which are an important additional input into the evaluation process. At the bases of many of these concerns have been questions raised by some students concerning the ultimate nature of the College's program. Is James Madison to be an *experimental* college in the sense of establishing a carefully designed innovative program of undergraduate social science instruction based on a professionally integrated multidisciplinary curriculum? Or, alternatively, should it become an *experimenting* college in the sense that educational objectives are loosely defined to enable individual students to have the maximum of educational freedom without the constraint of distributive and other requirements which seek to satisfy specified objectives of a liberal arts-social science program? Much of the discussion in the College this year over the relevance of particular requirements has stemmed from general student demands for more innovation of the "experimenting" variety. A review of the relative appropriateness of these various educational models and the need to consider to what extent they are mutually inconsistent or capable of integration will necessarily be a major concern in the College next year.

In addition to these philosophic issues, another ingredient of the evaluation process concerns the expectations of students enrolling in James Madison College. Appropriately, students have high expectations about the quality of the Madison program and this setting of high standards is entirely justified based on the College's commitment to undergraduate education. However, many students may have adopted an extremely idealized view of what can be properly expected as "excellent" instruction based either on an inappropriate "spell-binder" model of effective teaching or an unwillingness to apply the same high standards in judging undergraduate instruction in the wider University as well as in Madison. These concerns, together with a recognition of the faculty's deep devotion to undergraduate teaching as measured by the long hours spent in committee meetings and on individual class preparations, pose a real problem for the Madison community. How can the College live up to its general commitment to excel in the quality of its instructional programs while at the same time being able to properly meet student expectations of what can be expected within a context of limited human and physical resources? Balancing this ideal with reality is a major challenge to Madison's academic program.

The Planned Fifth-Year Evaluation

Major efforts during the next academic year clearly must be devoted to readying the College for its general evaluation. Extensive data for this evaluation have already been collected and the College has been careful to prepare extensive annual reports to provide a documentary record which ultimately can be used for evaluation purposes. However, more work needs to be done on a clarification of the College's goals to enable them to be operationalized for evaluation purposes.

Specifically, next year a systematic collection of the following kinds of data in preparation for the fifth-year review is being planned:

1. *Objective measures of academic achievement.* Graduate school and/or job placement experience of Madison graduates; evidence on the relative performance of Madison students in courses offered elsewhere in the University as well as on the Graduate Record Examination, Miller Analogies, or similar measures for graduate school entrance.

2. *Subjective measures of academic achievement.* An evaluation of the adequacy of the College's social science curriculum by an outside board of distinguished educators and scholars as well as subjective judgments about the reading, writing, reasoning, and calculating skills of Madison students based on their experiences with the College's annual writing requirement, the social science methods program, cocurricular programs, and the Field Experience program.

3. *Impact of Madison's special living-learning environment.* Evaluation of its impact on the scholarly values shared by the Madison community as well as on the personal growth and development of students.

4. *Relevance of general behavioral science studies of the impact of environment and curriculum on the learning process.* Evaluation and possible application of appropriate research studies; e.g., the research of Prof. Theodore Newcomb of the University of Michigan to the Madison experience.

5. *Costs and resource allocation.* Evaluation of the internal costs and benefits of resource allocation in the College as well as continuing efforts to devise an effective basis for cost comparisons with other units in the University.

Finally, the College reaffirms the need for the involvement of an outside board of distinguished educators of impeccable integrity and competency in any general evaluation of the College's program. This evaluation model is not only supported for its own intrinsic merits but also because of the College's belief that the ultimate decision as to the value of the Madison program must strongly reflect professional judgments and not merely the existence of intrainstitutional struggles based on specific political biases or economic rivalries for the University's scarce resources. Thus, in addition to the evaluation efforts supported by its own staff next year, the College plans to continue its search for foundation support to finance such an independent review body. It is hoped that this panel can be named and can begin its preparation for our fifth-year evaluation during the coming academic year.

Chapter 11

WE KNOW THAT YOU'RE OUT THERE, BUT WHAT ARE YOU DOING?

Larry H. Litten

We will begin with the assumption that evaluation is intrinsic to action. Paul Dressel noted this at least 10 years ago when he stated, "When one is faced with choice, evaluation, whether conscious or not, is present."[1] The foundation of these new colleges whose reports we are considering involved evaluation, as does their continuation. Again, Paul Dressel: "The only issue is the manner in which the evaluation is done and the extent and type of evidence used."[2]

Actually, we are dealing with three phenomena: Colleges as educational institutions, activities that collect information about colleges, and reports that communicate information about colleges to us as outsiders. What we can know about colleges as educational institutions is limited by the nature of the evaluation activities that take place with respect to the college. Both what we know about the colleges and the evaluation activities are further limited by the quality of the evaluation reports that are generated. We should be able to expect an evaluation report to tell us the following things:

1. What the participants in an educational program were trying to do.

---

[1] Paul L. Dressel, "The Essential Nature of Evaluation," in Paul L. Dressel and Associates, *Evaluation in Higher Education* (Boston: Houghton Mifflin, 1961), as cited by Lloyd J. Ring, Chapter 9, this monograph, p. 188.

[2] Paul L. Dressel, Chapter 1, this monograph, p. 1.

2. How they tried to do it.

3. How things worked out, with some evidence that things are as they say.

4. How the evidence was obtained.

Arthur Chickering has set forth five excellent principles for evaluation data.[3] To them I would add only the following: that the data be useful in making evaluation reports and that they be used for this purpose. I would like to address some specific comments to each of the reports and then attempt to draw some conclusions concerning where we are with respect to evaluation and evaluation reporting in the new colleges.

The Johnston Report[4]

This report specifies a large number of institutional objectives. It talks about some types of development sought for its students and identifies a number of activities intended to promote this development. An extensive program of systematic data collection on its students and a comparison group is described in a sketchy fashion. Considerably more attention is given to less public forms of evaluation.

Youngest of the colleges reporting, evaluation has already taken place at Johnston according to the report. Appended to it are the evaluative comments of several of the college's faculty members. Reference is made to the fact that "since the data indicated a considerable measure of success in realizing learning objectives, all [faculty] were reappointed." At another point, it is reported that some of the most exciting learning has taken place in seminars devoted to faculty research interests. There is also a suggestion that findings have resulted from interviews conducted by an external agent. However, none of the specific data alluded to is reported.

Many of the objectives stated for the college are what Arthur Chickering calls "hopes for institutional practice,"[5] and they are very elusive. However, this report is one of the most faithful in linking student developmental objectives with at least minimal indication of the means chosen to facilitate their achievement. Some of these objectives are quite specific such as the requirement for interdisciplinary studies that "naturally involve a grounding in the fundamental fields of mathematics, physics, chemistry . . . and communications." However, most of the objectives are highly abstract (e.g., "to help each student realize his full potential as a person, with a continuing concern for the determination of those values which form the basis for attitudes and decision making."). It makes me particularly uneasy when a faculty member suggests that the college is concerned with phenomena that are "subjective par excellence." How will the relative value of the enterprise ever be judged on those terms?

---

[3] Arthur W. Chickering, Chapter 2, this monograph, p. 27.

[4] Chapter 3, this monograph.

[5] *Op. cit.*, Chapter 2, p. 29.

There appears to be a considerable amount of elusive expectation lurking behind a cloak of libertarianism at Johnston. It will be interesting to see where and how the balance is effected. Built on top of the assertion that beginning students are "simply asked, 'What is it you want to do?' " and told to "begin where you are and work out from there," there is an elaborate structure of amorphous expectations contained in the guidelines for individual learning contracts. These are developmental objectives that reflect the noblest of liberal arts concerns in the loosest of liberal arts language, frequently with some of the more contemporary flourishes. There is little suggestion of what the criteria of achievement might be. It is implied that they will be worked out ad hoc between the student and his committee. But the only guaranteed quality control is the intellectual integrity and competence of those immediately involved in the contract evaluation. This is no worse than traditional methods of monitoring development and certifying competence, but it does not seem to be as great an improvement as we might hope for.

There is an unfortunate divorce between the evaluation of individuals on the basis of the contract guidelines and the evaluation of the institution as a whole. The listing of baseline data-gathering instruments for assessing changes in students ignores actual performance data for the most part and relies primarily on self-report instruments. Interviews by an external agent are one of the most promising components in the evaluation package and one of the best evaluation devices mentioned in these papers. Hopefully these outsiders also will look closely at the early and later work produced by these students and observe them over time as they engage in their learning activities. The interviews themselves can be a fine occasion for obtaining direct evidence of intellectual and interpersonal skills. Group interviews can be particularly productive. But the regular curricular situations in which students' intellectual and interpersonal skills are challenged and rehearsed can be the most revealing source of developmental data. Either through direct observation or via mechanical recording devices, such data should be gathered.

Interviews, and hopefully observation, by an external agent promise some important evaluative feedback for the college. It should be a regular part of evaluation and such evaluation should be a regular part of a college's existence. But how long can a college afford such a program? Are there ways that it could be institutionalized? Could students from other colleges do much of the work for their own educational benefit and to the advantage of the college? Perhaps existing associations in higher education could coordinate an intercollegiate evaluation exchange, hopefully with training for participants and quality controls built in. Perhaps consortia specifically for this purpose should be developed.

Many of the most interesting questions raised by an institution that is as far to one end of the self-directed institutionally guided continuum and the interpersonalist-intellectualist continuum as Johnston seems for the most part to be, will not be addressed by the evaluation program as described. Questions for which we would like at least tentative data include: How do persons (and which types) from institutions at various points along the continua compare in the breadth and depth of their knowledge and their competence in pursuing an inquiry and applying knowledge and methods from various spheres to problems of diverse sorts? How do they compare in the breadth and depth to which their values and attitudes have been tested against alternatives and in their development of self-satisfying and socially productive values and attitudes (including moderate dissatisfaction)? How do they operate in situations where various interests are in conflict, either for the individual or between persons? How well can they engage in criticism, experimentation, and evaluation?

Johnston's institutional evaluation design can within the limitations of the instruments used provide some evidence of its effectiveness relative to one alternative institutional arrangement, its parent institution. Again, it would be good if the design could include observation of students from the two colleges engaged in parallel learning tasks and even working together on the same task. Given the apparent distinctiveness of much of the Johnston approach, however, it would be even more desirable if the evaluation could be expanded to embrace comparison with a number of institutional models. This will require, at the least, evaluation reports from other colleges and from Johnston that would permit such comparison. Probably it will necessitate a much more intentional effort to develop intercollegiate evaluation activities.

Johnston must be commended for its explicit concern with faculty evaluation, even though the methods for effecting it are unclear. Hopefully it will be tied in with a faculty development program. Only two of the reports specifically make note of this important aspect of institutional evaluation.

The Morrill Report

D. Gordon Rohman's report on Justin Morrill College (JMC) at Michigan State University is certainly one of the most provocative papers in the group.[6] Although most of the papers mix "hopes for institutional practice" with objectives for student development, JMC, as reported by its Dean, seems to see institutional practice as its primary objective. The central question addressed by the author is, "What is the function of our college?" The answer comes not so much in terms of facilitating the development of its own students, as calling into question all established educational practices.

For me, the first issue is "how does JMC go about trying to perform this function?" A couple of examples are given and hints dropped about its general practice: JMC tries to do things differently than they are done in other American colleges. However, if the contents of the report reveal anything about what goes on at the college, it would appear that a good portion of the time is spent extolling the fundamental desirability and exhilaration of change and suggesting that anything that is being done, at JMC or elsewhere, may well be unnecessary and undesirable. Unfortunately, little attention is given in the report to what is being done and why, or to consideration of the rationale for specific changes and the determination of their consequences.

Mr. Rohman suggests that the best evaluative questions that can be asked of the new colleges at this point is, "What questions are you asking?" and "What are you unlearning?" He may be right in saying that negative reactions are the first step toward learning for the undergraduate. The college should serve these undergraduates, but need it behave like them (or a caricature of them) in order to do so? Is there not an important distinction between reflexive rejection of what is, and questioning? And is there not an even more important one between questioning, and inquiry or experimentation?

Mr. Rohman suggested several interesting scales on which an institution with the radical objectives of JMC might be assessed. Some of them might work.

---

[6] Chapter 6, "Evaluating Change in New Colleges," this monograph.

But first we would require an evaluation report that provides some information appropriate to the scales. One scale is the Risk Scale: asking how much the college "bit off" not how much it "chewed." If there were any standards by which such a scale could be devised, one would want it also to answer questions like "When did the college begin to choke?" or "When did indigestion set in from improperly chewed ingestion?" There is the Shoreline Scale (How far out have we waded?) which should be supplemented with a Swimming-Floundering-Drowning Scale. I like the idea of the Question Scale and wish that Mr. Rohman had been more specific with respect to the questions that JMC is asking. However, in addition to knowing how many questions were asked, I'd like to be able to determine how well the questions were formulated, how extensive was the the information brought to bear upon them, and how soundly was the evidence weighed. Most of all, I like the Wastebasket Scale ("what we have found out that won't work and *why*"–italics is mine), because it might prove instructive to people who read evaluation reports. I'd like to see the Wastebasket Scale become a Wastebasket Report and see it complemented by a Top Drawer Report (what worked well and with whom). Above all, I'd like the data from which to make my own evaluation.

The role of educational critic that the college has assumed is elaborated by reference to two crises that face institutions involved in transformative change: the crisis of unlearning the past and the crisis of inventing alternative futures. I would like to add a third "crisis," although I prefer the term "task." That is the necessity of learning from the development of alternatives, of "mapping" these uncharted waters into which, as Mr. Rohman claims, the transformative institution is presumably sailing. To confront the question, "How do we know and communicate what we are doing?" is probably one of the most radical challenges that an institution of higher education can accept.

Thus, one of the most important evaluative questions that can be put to these new colleges is, "How good are your evaluation reports?" How well have you done what we should be able to expect of an evaluation report? (Of course, the question becomes meaningless if the college has been so transformative that its question has been "Why should an educational institution facilitate learning?") It is difficult to accept the claim that it is too early to look for results. Certainly we will need longitudinal research on both individuals and institutions in order to get very complete answers. Nevertheless, consequences must certainly be in evidence after four years of "sailing"–data that could provide us with invaluable insights when we get to the longitudinal information. Political realities cannot be ignored. However, I feel that it makes good politics to be able to say, "This is what we have discovered so far and these are the reasons why we feel that the data are inconclusive."

Man cannot live by questions alone and Morrill is more than an aggregation of questions. It is a number of tentative answers. The report mentions starting down one road and then taking off on another. Surely evaluation is implicit in such choices. The paper talks about the use of a number of instruments to obtain systematic feedback on its students at various points in time and claims that these devices have provided "important and valuable insight about our various programs." But we are not told the nature of these insights nor given the data from which to derive our own. Mr. Rohman claims that it is inappropriate in an evaluation report "to detail the number and kinds of changes we have made at JMC in the past five years." I would submit that that is precisely what should be done along with indication of the information that led to both changes and continuations. Other institutions might well be alerted to possible problems and promising alternatives.

The educational objectives claimed for JMC give us an interesting instance of what Paul Dressel refers to as the tendency in innovators to consider their institutions and institutional objectives to be unique. We are told about four "attitudinal competencies" which the college seeks for its students. First, it should be noted that a college presumably dedicated to unlearning should be cautious not to fill the vacuum that it creates with jargon (e.g., attitudinal competencies) that will require additional unlearning. We are presented with what are essentially two attitudes (open-mindedness and a tolerance for ambiguity), a competency (the capacity to ask "transformative questions"), and something that appears to be a variety of competencies (extended modes of awareness beyond the cognitive, problem solving, and system building). What is intriguing is the assertion that these objectives are "relatively uncommon in lists of traditional academic learning," but "are nevertheless more and more frequently being cited by those who study the future as being relevant skills for coping with change." Even without resorting to the dubious authority of those who "study" the future, it is remarkable to note the familiarity of these objectives in liberalartsese. Except for the last objective, whose meaning is unclear, and if we translate "the ability to ask transformative questions" as "critical thinking," I think that few persons in academia would be uncomfortable with these educational goals (even though many might not be very good at assisting undergraduates to achieve them and might be extremely up tight about actually having to do their utmost to try).

Mr. Rohman draws a fairly sharp distinction between acting and thinking, and suggests a corresponding division of labor between the innovator and the evaluator. Certainly independent evaluation is to be desired and sought as an intidote to the myopia that frequently develops in the course of any consuming action. But the actor cannot place total responsibility for feedback on the evaluator. For one thing, evaluation needs to be too continuous and comprehensive to permit that to happen. For another, the innovator can clarify his reasons for action, heighten his awareness of the effects of his actions, and contribute to the enlightenment of others by attending directly to the problems of evaluation. To talk as though actions had no discernible effects is ludicrous. To place total responsibility for the determination of those effects with others is irresponsible. No explorer sails off the maps without some instruments to guide him and without attempting to map his discoveries. Mr. Rohman recognizes the need and indeed welcomes evaluation. But the college, new or old, holds most of the responsibility for evaluation—a responsibility primarily to its students, but also to the society it serves and to other colleges. The new college has the additional opportunity to inform others of ways of doing things better or ways of doing better things.

The Fairhaven Report[7]

One of the finest aspects of this paper is its recognition that there are multiple functions that the collection of evaluative data should serve and several audiences to which it should be useful. The pervasive tone of the paper is that "we need to know what the effects of association with our institution are and we need to be able to tell others about them." This is an exemplary stance toward evaluation.

The three strategies proposed actually consist of two strategies (the collection of data within the college and the longitudinal studies of alumni), two techniques for data gathering ("unobtrusive" systematic observation and standardized instruments), and a principle to be followed in every strategy (translatability of the findings into laymen's language).

The follow-up strategy and the systematic observation technique can be very expensive if done right and may be feasible only if they are used as educational activities in themselves. Both of these devices could easily be used to this advantage. When such studies are designed and executed as part of a course on higher education or related problems in the social sciences (particularly social research methodology or the evaluation of social institutions), or as part of an independent study project, costs can be reduced by making learning one of the principal payoffs for engaging in the activity. In addition, the work can be both useful and educationally challenging. The more intercollegiate such activities were, the stronger they could be as evaluation devices (cf. above, comments on the Johnston report).

---

[7] Editor's note: Mr. Litten did not see the complete paper on Fairhaven College in Chapter 4, specifically the closing remarks by Paul Woodring. Mr. Woodring added the concluding remarks after the death of the author, Dean Charles Harwood.

The concern with making the results of evaluation intelligible to the lay public is highly commendable. It will require much educational effort, however. It would be nice to educate the public to look for institutional effectiveness in facilitating development instead of looking only at the outputs. But educators are still having enough difficulty talking among themselves about institutional effects and effectiveness.

It is interesting to note the observation from a recent conference that leaders in innovative higher education were uninterested in research reports on the personalities of college students. I was at a similar conference (or perhaps the same) and received the same impression, although it was partly a function of presentation and not substance. It is impossible to determine just what else Mr. Harwood's paper would have included. It is disturbing to note, however, that the evaluation program as far as it was described focuses on experiences that the student has and on personality change. There is almost no mention of intellectual competencies beyond subject-matter mastery.

The behavioral observation approach has tremendous appeal, but it raises some serious issues of privacy and "big-brotherism," particularly in situations that may not be as open as my impressions suggest that Fairhaven is. In such undertakings the subjects should be aware that they are being observed and should not feel threatened by such observation. At the same time, a case can be made that persons not willing to engage in such self-scrutiny or contribute to institutional improvement should be elsewhere than in college. In any case, confidentiality should be guaranteed individuals. Above all, the subjects should be involved in requesting and designing this type of evaluation and in determining the utilization of the data so obtained. It seems foreign to the ideals of higher education not to desire, in fact seek out criticism. But we must also be sensitive to the need for privacy. One would hope that the same careful scrutiny applied to the out-of-class behavior of students would be applied to the course-related performance of both students and professors. If the extracurricular behavior of students is a fair focus for evaluative data collection through observation, is that of professors equally fair game? If so, limits may be drawn rather quickly.

The attempt to exploit detailed faculty evaluations of student development by setting some guidelines is particularly promising. The use of outsiders to examine student work and observe behavior would be a fine complement and quality control to these procedures.

Finally, it is encouraging to note an approach to evaluation that sees it as intrinsic to institutional development and responsibility. Mr. Harwood says "The identification of college experiences both in and out of the classroom and the investigation of their relation to later life is the key task of the evaluator as well as the planner of higher education." The responsibility should be extended to include all participants in the educational enterprise.

The Michigan Report[8]

This is the best evaluation report in the group. Data collection was extensive. The design for their acquisition, while containing gaps (e.g., incomplete data on control groups), was well conceived. The analysis was highly sophisticated. But most importantly, we are presented with the data and a fairly good account of their acquisition. The analysis is performed on the data that are presented to us. Thus, we can make our own judgments and draw some comparison with our own frames of reference.

The findings generally support previous evidence that small, "innovative" units within larger institutions tend to attract distinctive students, lose a good portion of them, and that distinctions between this group of highly interacting persons and persons in other settings increase over time. Whether these things are desirable, as the report implies, is debatable. Priorities should be more with educational improvements instead of recruiting advantages.

One of the most interesting suggestions is that the Residential College (RC) serves a particular range of persons as identified by personality scales. Persons above or below this range tend to move on to other environments. This supports the hypothesis that has been advanced elsewhere, that given colleges most effectively serve particular types of students (Chickering & McCormick, 1970). Persons outside of a given college's range, particularly those who are above it, may find more reward in moving on to another environment. Certainly this becomes a very important question if evaluation is viewed as a means to deciding about the most effective allocation of scarce resources. A good evaluation program in a college should help to determine when a student might most productively move on to other activities or environments and be able to report data on institutional effects and effectiveness derived through such a framework.

---

[8] Theodore M. Newcomb et al., "The University of Michigan's Residential College," Chapter 5, this monograph.

To me, the data imply that more intellectually oriented students are attracted to the RC for affiliative reasons. The authors claim that lack of difference on the OPI sociability scale is evidence that the students are not seeking such ends in coming to the RC, but turn around to report that where there is a significant difference between the control group and the RC students on the CUES scales, it is in anticipations of close community and faculty-student relationships.

There are two problems with the control group comparisons. First, it is not clear whether persons who dropped out were removed from the statistics on entering students. Secondly, one might find a more informative pattern of effects if the larger, heterogeneous LSA group could be disaggregated into subgroups with more likelihood of high social interaction (e.g., other living groups) or different amounts of exposure to RC-type courses (e.g., various majors). Some groups might look very much like the RC.

With these qualifications, the supporting evidence for the accentuation theory raises some intriguing problems of educational philosophy. Is it desirable from a larger social perspective to create quality institutions through the attraction of distinctive students at higher states of development and then accentuate differences? Does it have a deadening effect on the institutions from which they are diverted, or might they be lost to any other institution of higher education? Does it promote undesirable forms of separatism and elitism? We need more behavioral data that can be brought to bear on these questions.

Another question that the findings suggest concerns the nature of what actually happens in the changes that we observe. The scales are mostly self-reported attitudes (likes-dislikes, opinions). They do not include behavioral evidence of these attitudes. The question then becomes, how much of this change might be mastery of institutional rhetoric? How much might be conformity to social pressures, including the pressure to claim nonconformity? We have no evidence to suggest that either is true, although other researchers working in college evaluation recently have questioned the validity of change so reported (Suczek & Alfert, 1970). What we would like is behavioral evidence to the contrary. Inventories and questionnaires may identify some areas for further exploration. But they do not reflect the extent to which the reported attitudes and values have been examined critically. They certainly do not demonstrate either competence or development in the ability to use knowledge and intellectual skills or to work resourcefully and effectively toward socially responsible ends. We would like performance data that show an increase in an individual's options through the development of awareness and competence.

In spite of its good presentation and analysis of data, as an evaluation report this paper has flaws. The most annoying is the well-stated argument, albeit mostly by innuendo through reference to "typical" evaluation reports, concerning the need to specify institutional objectives and to measure them with carefully selected instruments. This is followed by the pious assertion that this is what these evaluators did. We are then referred to an appendix for a list of these objectives, but find nothing more than a superficial description of the structure of the college and the specification of a few formal requirements (e.g., foreign language proficiency). It is true that the authors make oblique reference to a couple of objectives in the discussion of the findings. They state that "a sense of community and campus morale–a goal of the RC planners–has apparently developed. . . ." Further on, they suggest that "it may be a source of some disappointment to the RC planners to note that intellectual commitment or awareness are no more emphasized in the RC image than in that held for LSA."

There are other omissions such as the failure to specify use of an atypical scoring method for the CUES instrument, failure to report the wording of the instructions used with CUES (each group was asked to respond in terms of the RC or LSA, not the University), and the previously mentioned failure to indicate handling of the dropouts. But all in all, the Michigan report is a good one that at least allows us to sharpen our questions. Evaluation data appear to have served this purpose internally in other institutions (e.g., Callison), but the Michigan report assumes a much broader and commendable responsibility.

The Callison Report[9]

As a report on evaluation activities, the Callison paper is among the best received. Educational objectives are specified, although they can hardly be considered ends that Callison as a college self-consciously purports to serve. The report of the enterprise required to dig out these objectives from a variety of sources is interesting acknowledgment of what the other papers state by omission: colleges are not easily pinned down into saying exactly what they will try to do for people who give them their time and money.

---

[9] Beth B. Mason and Douglas R. Moore, "Toward a Strategy of Evaluation for a New College," Chapter 7, this monograph.

The paper gives a general description of the structural and programmatic features of the college, evaluative data-gathering devices are identified, some data are provided (although casually and mostly on entrants), and some extremely provocative questions are raised concerning the nature and effects of such an institution. This report is the only one that identifies specific payoffs from the collection of systematic data in making changes in the institution and in anticipating problems, and at the same time provides us with the data. However, there are also many general references to changes made at Callison or problems foreseen without provision of any supporting evidence.

One of the most notable aspects of the paper is its explicit recognition of how informal and formal systematic evaluation complement each other and how the former should provide leads into developing and improving the latter. A fine example is the discussion of the anticipated movement from close personal observation of students by the faculty to the more efficient and systematic use of objective data in determining readiness for the India program. This type of evaluation represents the orientation that I find most appealing: determination of when a student is ready to move on to other learning activities. Callison seems to have achieved an unusual integration of individual and institutional evaluation. If this kind of data is systematically collected, it can then be used for institutional evaluation in terms of both student change and institutional efficacy in the promotion of development. Of course, such data must be supplemented by other information on institutional practices and their consequences.

Callison's involvement of students in the collection and consideration of evaluation data is exemplary. Students should receive such feedback and be able to use it for their own development. Of course, this requires evaluative data that represent actual skills and attitudes, not learning the acceptable answers. The practice of student generated and executed systematic evaluation research is particularly worthy of note.

It is unfortunate that the Callison evaluation activities do not have a broader perspective, and seem to be retreating further from the question of institutional effectiveness in a comparative framework. The degree to which mere institutional visibility is an unsatisfactory evaluation device is emphasized in this paper by reference to the lack of feedback from outside researchers. I will not go to any length to defend the unresponsiveness and irresponsibility of such persons. However, I think that the remarks of the student about how it "feels like being in a zoo," and the complaint concerning the lack of feedback may reflect an unfortunate but not uncommon situation. Often much more is expected from external observers than the present casual practice of the art permits it to deliver. Where some systematic data gathering has been attempted, some formal feedback is the least that can be expected. But many visitors are simply trying to begin to get a hold of something and make some very crude personal assessments. Openness to such curiosity is part of the price to be paid for claiming to do something different or better. The more evaluation reports reflect what actually happens in these colleges and how, with data that can be believed and used by outsiders, the less intrusive such visits need to be.

The questions raised in this paper for further investigation on the small, distinctive college are extremely important ones (e.g., the possibility in these colleges of concern for the affective intruding upon and crowding out intellectual development instead of complementing it, the potential problems of conformity, elitism, and escapism). Unfortunately, there appears to be little in Callison's formal evaluation program as described that is directly addressed to these questions. Mention is made of the fact that "casual observation and our more systematic evaluation have led us to conclude that a cluster college is particularly vulnerable to faculty members who need to collect disciples around them and to students who are seeking a 'guru-type' master." But what is the evidence? These are sensitive data, but having such information is essential in evaluating the risks and rewards of collegiate innovation in directions taken by the new colleges. This underscores a dilemma in public evaluation: outsiders need to have credible data on which to base their evaluations, while individuals and institutions need to be protected from inaccurate or incomplete data.

At one point, reference is made to persons who "prosper" in the Callison setting versus those who are negatively affected. The data sound fascinating, but "prospering" and "negative effects" are never defined nor demonstrated.

An extremely interesting point mentioned in the paper is the usefulness of lessons learned from the other cluster colleges at the University. We would like to know the nature of these lessons, the type of information included in them, and the form in which it was obtained. I wonder what lessons would be learned from evaluation reports like most of those that were obtained in this effort.

The concern about the adequacy of traditional data and criteria for judging places like Callison is common among these institutions. However, as is also common, little is said about what these traditional data and criteria are or why they are inadequate. Those that are identified in the report are not among the more noble objectives of the traditional liberal arts college. If, as suggested at one point, the "effects" of Callison may be due simply to developmental processes in youth, we can call into question the need for Callison and other colleges and begin to explore other uses of these resources. The ultimate criteria suggested, satisfaction and fulfillment in later life, while undoubtedly of great importance, cannot stand alone. Given some of the characteristics of entering students that this paper reports and some of the potential problems in such colleges that are identified, self-satisfaction grounded in intellectual competence and an understanding of others seems like a more satisfactory goal. And perhaps a touch of dissatisfaction should be there too?

The New College Report[10]

Along with the authors of the Callison report, Mr. Elmendorf has identified very explicitly what most of the other papers have ignored: the complementary role of objective, extrainstitutional indicators of institutional effects and the observations and deliberations of intelligent, sensitive men within the institution. The objective data cited are unusual (e.g., publications) and rest on a sound principle: the judgments of intelligent men obtained from a broad base beyond the college. However, there are no control or baseline measures for such data. Thus, there is no indication of actual intellectual development. It may well be that New College affords an occasion for persons to perform work that gains recognition from the intellectual community outside of the institution, but there is nothing that says people *learn* to engage in inquiry or communicate effectively as a result of the New College curriculum. There is nothing that says it happens as well or better at New College than might be expected elsewhere, given the same caliber of student. Furthermore, for how many persons, even at New College, are publications per se a useful indicator of competence? Certainly the ideal of the paper well enough thought out and written as to be suitable for publication is a fine one. Can New College bring all or most of its students to such competence? How much of this can be attributed to the curriculum or structure of the college?

It is unclear exactly how these external data are used in summative evaluation of students. They are an interesting form of quality control on the faculty's judgments. It would be nice to see it done on a more regular basis through intercollegiate cooperation in systematic evaluation work that would include the "referring" of papers written at various stages of a student's academic carreer. Paraprofessionals could contribute much to this function both within and across institutions. Such review should also be extended to other competencies.

---

[10] Chapter 8, this monograph.

Mr. Elmendorf refers to the "new values" that these new colleges "beget" that demand new measuring devices. What are these values? How is New College attempting to promote these new values? And how is it determining its effectiveness? An evaluation report should tell us all of these things.

The role of the College Examiner at New College has always intrigued me. I am surprised to learn that the functions he performed were found to be limited in their usefulness and that certain aspects of his role could not be expanded to maintain the usefulness that the office had during the early years of the college. I regret that the paper does not answer the questions that the decision to abandon this role raises for me.[11] Why do comparative studies have to be grounded in "false" norms? What prevents an examiner from developing new measuring devices appropriate to the "new value systems" of new colleges? Could a college examiner work cooperatively with several colleges so that costs could be shared and more valid comparative data could be generated? As other colleges continue to innovate or wish to do so, why doesn't New College wish to compare the effectiveness of its structures and programs with that of other colleges, either as a means to self-improvement or to alert others to potentially useful devices for both education and evaluation?

Mr. Elmendorf warns against the "trap" of seeing one's problems as unique. In reading these papers, it is apparent that the handling of these problems is generally not unique either. A commendable self-critical stance is claimed by most of these institutions. However, the inability to communicate the data upon which such evaluation is based is a common problem.

The University of California at Santa Cruz Report[12]

This is the only report that admits that systematic formal evaluation is essential in colleges and at the same time openly confesses to widespread resistance to it. It is regrettable that a major new institution like the University of California at Santa Cruz (UCSC) has not seen formal evaluation as part of its contribution to American higher education. The misfortune is compounded by the fact that the basic cluster college structure of the institution affords some unusual opportunities for research in higher education. This paper underscores the important fact that systematic evaluation that is not built in at the beginning of a program will probably be achieved only with difficulty, if at all.

---

[11] AT the AAHE convention, Mr. Elmendorf suggested that there is a newly felt demand at New College for a less test-oriented examiner.

[12] Lloyd J. Ring, Chapter 9, this monograph.

It is important to note the positive reception given the research reported in this paper in an environment as resistant to evaluation as Santa Cruz appears to have been. It speaks well for either the evaluation model adopted by the evaluation group at UCSC or their human relations skills or both. It is a good example of institutional research serving as a means of calling attention to particular issues, a spin-off emphasized by Chickering.

The research reported for Santa Cruz portrays some interesting sociological aspects of the college. It is suggestive information for the administrator concerned with problems of organizational cohesion and accommodation between various groups within the organization. The data on living accommodations preferences should be of interest to institutional planners as well as administrators, particularly those who found their dreams on the living-learning group.

As reported in this paper, evaluation at Santa Cruz lacks a focus on the colleges as a means of fostering student development. The emphasis of the evaluation is on organizational-administrative matters, although it is acknowledged in the discussion of the student self-rating data that the lack of longitudinal data is unfortunate. Hopefully, the ambitious and conceptually admirable plans for an evaluation program will take a turn toward the student and concern for the optimization of his educational development. It would be nice if such a concern could focus on demonstrable competencies and not just self-reported attitudes and self-perceptions. The adoption of Chickering's data-gathering design within the framework of the evaluation model Mr. Ring describes would be good. Application of the resources and research talents at UCSC to the development of better behavioral indicators for use within such an evaluation model, particularly if the information is useful for individual student guidance, would be a major advance toward institutional responsibility in higher education.

The James Madison Report[13]

In a certain sense, the James Madison report says a good deal. It tells about the structure and programmatic features of the college. It mentions a number of educational objectives and ties them to particular means of achieving them, mostly courses. It suggests some of the activities that occupy the time and result from the initiative of students–things that sound like acitivites where intellectually alive students might well expand their awareness and develop their competencies. There is mention of a variety of evaluation devices. Reference is made to one of the most promising (but potentially useless) evaluative devices specified in these reports: the evaluation of the curriculum by an external group of "distinguished educators and scholars."

At the same time, there is much lacking in the paper as an evaluation report. It alludes to data that have been collected but doesn't present any. Likewise there are a number of evaluative comments (e.g., the field program was an "unqualified success," several suggestions of the need to revise something) without supportive evidence. Having been supplied with specific educational objectives and the activities chosen to facilitate their achievement, we are not provided with evidence of how well and for whom the objectives have been achieved.

There seem to be some unjustified vestiges from traditional curricula in the Madison program. An example comes through in the discussion of the Foreign Language Requirement and what are referred to as "substitute liberal arts cognates." There is no rationale presented for considering the various alternatives as equivalent or for making any of them requirements. These are very different activities. Are they really substitutes? Here is where the transformative questions that Madison's sister institution (Morrill) claims to be asking should come in. One would ask, "What are language requirements supposed to do and is it worth doing?" "How well are they succeeding?" "Are there ways to make them more effective?" "Are there other ways to do the same thing?" Information should then be gathered on the alternatives, either from other persons' experience or through experiment or both. Presumably, these questions were asked in the course of deliberations on the matter. However, if the change was significant to Madison's objectives, then the evidence on which it first became a question and on which the question was resolved should be part of the evaluation report.

---

[13] Herbert Garfinkel, Chapter 10, this monograph.

Part of the problem with the James Madison paper as an evaluation report derives from the fact that it is a modified annual report for internal public relations. While it is not bad as the genre goes (although it would be hoped that evidence of actual learning and student development would become a regular feature of such reports), it is filled with a good deal of administrative detail that has little payoff for us. An evaluation of students' development that consists of reporting the completion of various program requirements tells us little, particularly if we are trying to see the particular institutional arrangement as an alternative to other means of facilitating development.

Madison's plan for evaluation is relatively broad in its expressed concerns but not very thoroughly presented. It seems to lose sight of many of the intellectual development objectives mentioned in connection with the courses and some of the other aspects of the program. The evaluative data would be much enhanced by focus on student development over time. The basis for securing such information is not evident, except perhaps through the annual writing requirement. Proper exploitation of the latter would be indeed a good basis for evaluation at both the individual and institutional level. Although not irrelevant, the use of traditional academic output measures seems unfortunately restricted.

All this must be said with some caution since we are not given a very clear picture of how the "outside evaluators" will operate, what kinds of data they will have, and who will be making the "subjective judgment" of the intellectual skills that are made evident through the students' participation in Madison's curricular activities. The use of outside evaluators seems a superb idea, although one hopes that they will not just review the curriculum (the sentence is ambiguous). In addition to "distinguished educators and scholars," students who have flourished in other college programs should be considered as members of such a team, or comprise a separate one. Hopefully, these persons will not engage in the superficial attention paid to the trappings of a collegiate program that has characterized much past accreditation evaluation. Instead of looking at the curriculum, they should concentrate on behavior that the curricular activities evoke and changes that take place in such behavior over time. If these changes and the nature of the behavior can be shown to be distinctive from what is found in other programs, then it would be hoped that the key contributing elements in Madison's curriculum could be identified.

It is interesting to note that the Madison report is the only one that accepts the idea of evaluating a program in the framework of scarce resource allocation (i.e., costs and benefits). Perhaps that is due to its original audience. Nevertheless, it is desirable that institutions like Madison will contribute to the development of much more adequate benefit indicators and then proceed to face this vital question of costs squarely.

The Madison report alludes to the evaluation of faculty and suggests that an indicator of their quality is the number of invitations received to teach elsewhere in the university. One would wish for much better indicators of faculty performance than this.

Research for Action[14]

While not in any way an evaluation report, this is the most cogent argument for evaluation at the institutional level and the best explication of the basics in methodology that has come to my attention. The proposed design, with its several samples and variety of data-gathering techniques, is scientifically sound, aesthetically pleasing, and highly practical. The need for a variety of data-gathering devices that focus on both groups and individuals is important counsel.

The recommendations have several shortcomings. First, the evaluation devices suggested are generally too divorced from the actual educational activities through which students develop and rehearse their competencies and attitudes. Chickering's own *Experience of College Questionnaire* comes closest to tapping such information. But it is self-report and doesn't provide data on the development of intellectual competencies. The GRE will provide some information of use in this area, but it does not plumb the depth and cogency of students' thinking nor the skill with which they engage in learning activities. Even the discussion in the paper of case study interviews focuses on experiences, activities, and attitudes.

Although the principle is rightfully held high, little attention is given to the means by which institutional evaluation data serve the self-examination and education of individuals, particularly students. It should not be merely through identifying important developmental concerns by talking about them in questionnaires, nor giving students an opportunity to compare their own profiles with those of other students. The integration of institutional evaluation with individual developmental guidance by both exploiting learning situations directly for data on institutional effects and feeding standardized test and questionnaire information into the advising of individual students can enhance economy, validity, and usefulness of the data, and probably the willingness to cooperate.

---

[14] Arthur W. Chickering, Chapter 2, this monograph.

It should be noted here that there is a certain dilemma posed in the case study method. Paul Dressel suggests that it tends to cause the researcher to lose his professional distance from the subject, therefore compromising its research purity.[15] Arthur Chickering notes that for research purposes, the sample should be small or the data will be overwhelming. However, granting these research problems, we must remember we are dealing with educational institutions. Other researchers who have conducted studies of college students using intensive interviews, both individual and group, report that their subjects claim that the experience was one of the most rewarding educationally that they had (Suczek & Alfert, 1970, p. 138).[16] All students deserve these benefits. Conscientious advising, apparently something that occurs with considerable frequency in these new colleges, may accomplish the same ends. It is the least we can do for students. However, the research aspect may contribute in itself to the positive feelings of the students. Being party to a serious inquiry is rewarding. Our obligation is to make such inquiry serious. The best evidence of seriousness is use of the data to guide action. Again, intercollegiate exchanges for purposes of evaluation and advising through observation and interviews might help sustain this quality of seriousness beyond the regular advising interview.

The paper presents a very good design for the use of available instruments. Reliability is stressed. But reliability without the validity that we seek is a dubious asset. The possibility that findings are due to the recognition of "right answers," factual or attitudinal, must be reduced through data gathered from challenging situations. This is particularly true in colleges where institutional rhetoric about attitudes is laid on thickly. Data gathered from actual performance in educational activities would be an improvement, at least as a supplement to standardized instruments and interviews. Such evaluation should focus on behavior during and products resulting from these activities. This approach would benefit from institutionalizing evaluation at all levels. The regular recording of class sessions, particularly discussions, could provide such information for direct use by the participants and for other evaluators. Papers and other projects provide important developmental data over time. Admittedly, there will be reliability problems as we push into these areas. That's the challenge. We must try to bring reliability to the types of evaluation that reveal what a student knows and feels and how extensive the bases for that knowledge and those feelings are. We must try to do it in situations that actually challenge his knowledge, skills, and feelings. The instruments that we have available should be exploited for what they are worth. They should be recognized for what they can't do, as apparently they have been in several of these new colleges. The appropriate response to their limitations is to accept the challenge of finding the means to do the job more effectively.

---

[15] Dressel, op. cit.. this monograph, p. 14.

[16] Also see W. G. Perry, Jr., *Forms of Intellectual and Ethical Development in the College Years* (Cambridge, Mass.: Bureau of Study Counsel, Harvard University, 1968). Perry's work should be very suggestive in the development of analytic schemes for the evaluation of papers and discussions in many areas.

Conclusions: The Evaluation Reports as Reports

These reports from eight new colleges are a very diverse set of documents. Since the title of this volume is *The New College: Toward an Appraisal,* it must be admitted that we are a long way from the "objective evaluation of them" to which Paul Dressel refers in his introductory chapter. Some of the reports provide intriguing glimpses of new institutional arrangements for colleges; others raise provocative questions. Some do both. But we find out very little about what these colleges are trying to do, a little bit more about how they are going about doing things, almost nothing about what effects they are having, and only the most rudimentary idea of how they go about determining appropriate courses of action. On the whole, it appears that evaluation reporting is not a well-developed art. This in turn raises questions about how central systematic, public evaluation is in the concerns of these colleges. The ideal of the "examined life" seems to be part of a double standard–good for studens, but uncomfortable for colleges.

Most of the reports make at least oblique references to institutional objectives, but they are frequently the "heady" stuff to which Elmendorf refers or the "utopian and abstract" statements that Chickering mentions. Chickering's general observation on formally stated objectives holds true: "hopes for institutional practice" are frequently confused with statements of student development that the college purports to facilitate. The developmental objectives that are stated are rarely associated with specific devices for obtaining evaluation data or with criteria by which successful realization of the objective might be determined. Many of the papers refer to past or planned use of standardized tests and questionnaires. However, we are rarely told which of the college's purposes they are supposed to measure or why these instruments were selected. Neither are we told how they have been (or will be) administered or used. It seems that the mere listing of a string of familiar initials is supposed to suggest that the college is seriously undertaking its responsibilities to determine whether it is serving as more than an occupational and social class-sorting mechanism.

The papers contain uneven reports of the specific structures, activities, and personnel that comprise these new institutional means to facilitating student development. Nevertheless, there is much more talk about the way these colleges seek to do things than why or to what effect.

The reports talk a good deal about evaluation. According to the authors, in most of these colleges there is a good deal of institutional introspection. Mention is made of systematic data collection. There are references to the fact that changes have been made or practices continued as a result of evaluation. However, only three reports provide evidence from which we can draw some independent conclusions, and only one of these contains extensive data that deal with changes in students.

Conclusions: Types of Data and Data-Gathering Techniques

These papers report a wide variety of techniques for the collection of evaluative data. There are some very promising ones. However, apparently the most frequently used technique consists of the deliberations of persons within the college on the basis of their personal knowledge, experiences, and observations. Such deliberation is an essential part of institutional evaluation, and personal impressions are frequently the only readily accessible data. But the limitations of such information must be recognized. Impressionistic data generated by the participants limit the opportunity for external criticism and reduce the possibility of demonstrating to others the effectiveness of particular ways of doing things. Even intelligent men can develop biases, vested interests, and distorted perceptions. The very existence of the new colleges is testimony enough to their acknowledgment of that fact, at least as it pertains to the older colleges.

There is considerable emphasis on the students' perceptions of the college. This may help in institutional trouble-shooting, but it does not reflect student accomplishment or development. Personality and activities inventories are the most extensively used standardized devices. These rely on self-reported attitudes and activities. It is difficult to judge the validity of such reports so that even where data are presented, our ability to judge the effectiveness of these colleges is limited.[17] One wonders whether these students have actually changed in their values, attitudes, and behavior or have mastered the rhetoric of change, liberalism, and independence that seems to abound in these environments. Even if the data represent actual changes, we don't know whether they result from conformity to the prevailing norms in the college, or from a basic evaluation of the student's perspectives on the world in the light of other perspectives and the evidence supporting each. They tell us little about the students' ability to engage in critical thinking, cogent inquiry, effective deliberation of alternatives, the formulation of creative solutions to problems, or the ability to work productively alone or with others who may or may not share his values. Performance alone can tell us this with the confidence that we desire. Data from situations that actually challenge a student to act on the basis of his attitudes or to demonstrate his competence in the activities in which he claims to engage would be desirable, at least as supplementary information. Colleges use this type of situation to stimulate learning, and data from such situations are used internally to criticize and guide development. It is also on the basis of such performance data placed in a developmental perspective that colleges should be evaluated. If these new colleges are going to serve as means for introducing change into higher education, it would be nice if they did so on the basis of their demonstrated effectiveness in facilitating development in the types of students that they get, instead of on the basis of their aesthetic appeal as institutions or their ability to change questionnaire responses.

This brings us to the problem of effectiveness. The Michigan report is based on a design that had control groups and controlled for differences in these groups in the course of studying change. Other institutions suggest that they have done or will do the same. Most of them have the option of comparing their students with the normative groups of the standardized instruments. There seems to be substantial awareness of the need to control for starting points in looking at outcomes. However, there is little explicit concern for Arthur Chickering's proposition that "An educational institution should continue to exist if it fosters aspects of student development which would not otherwise occur, or would occur significantly less elsewhere."[18]

---

[17] It is interesting to note the seeds of disenchantment with personality change data in some of the reports, although they are still the mainstay of much evaluation. It is even more interesting to learn from the Johnston report (Chapter 3, this monograph) of the reservations that one of the developers of one of the most popular of these instruments has concerning its adequacy as a means of assessing what these new colleges are trying to do.

[18] Chickering, op. cit., this monograph, p. 30.

Comparisons, where they exist, are all within the framework of conventional academia. The means for other comparisons have not been readily available. However, the development of "open universities," "universities without walls," and external examinations systems could help us make these assessments. Hopefully, these new institutions will participate directly in this responsibility for more broadly based evaluation.

Even within the context of traditional academia, relative college effects and effectiveness as a problem of scarce resource allocation is not generally acknowledged. Only one report mentions costs as a significant criterion of institutional performance. A shortcoming of the reported institutional evaluation programs is their implicit acceptance of the traditional academic temporal lockstep. The focus is on what changes occur during fixed periods (e.g., courses, academic years, the four-year college experience). If individual and institutional objectives could be defined in terms of demonstrable competencies or attitudes, then learning activities and programs could be evaluated in terms of how efficiently they expand the options of individuals (or particular types of students) by bringing them to desired states and then guiding them on to other learning activities, or even other situations. Time would become an important variable in college evaluation, and the most productive use of an individual student's time a more central concern.

Some Further Steps in Institutional Evaluation

There is a significant gap between the evaluation activities described in these reports and the educational activities through which students expand their awareness, rehearse their competencies and the effects of their attitudes, and obtain feedback on their development from persons with more experience than they have. If the two levels of evaluation could be better integrated, gains would be made in validity, economy, flexibility, and cooperation. We need to balance individual challenge and the opportunity for originality with broadly based assessments of development within each student's program. Institutional evaluation also needs to reflect both concerns.

Cooperation between colleges could be increased not only to exploit learning resources (e.g., urban semesters, field stations), but also to expand the base of evaluation grounded in challenging learning situations. Programs where students come together from several colleges representing different institutional patterns could provide occasion for such evaluation (with the qualification that control for selective recruitment would be difficult). Intercollegiate cooperation in the evaluation of student work (papers, projects, etc.) or observations of educational activities (either directly or through video or audio recordings) could facilitate institutional evaluation where objective measures are difficult to obtain at present. Perhaps we could better exploit simulation games techniques or interactive computers to provide challenging learning situations with some constants that would provide the basis for systematic intergroup comparisons of competencies or behavioral styles.

Most basic, however, would be the creation of research and development arrangements in which these new colleges actively participated for the purpose of developing and effecting more adequate evaluation techniques. Faculty could be given leave to work in such programs. Since the evaluation of social institutions is one of the most challenging problems facing our society, students should participate directly in such inquiry. Exposure to the problems of educational evaluation, training in the available techniques, and an opportunity to do evaluative work would be good education. One of the best ways to get evaluation work done is through a regular seminar-workshop where students explore the questions and problems of higher education or the evaluation of institutional efficacy and also do some research on a particular question of concern to the local campus. The students can help define the problem, work out the information-gathering devices, collect and analyze the data, and write the report. Evaluation gains from student participation. Students gain from coming to grips with these very perplexing problems. Likewise, the use of student evaluation teams from work-study or independent study programs to collect and analyze evaluation data on campuses other than their own could result in substantial economies to colleges in doing the most expensive and labor-intensive types of evaluation work (e.g., participant observation, interviews), and important educational rewards to the participants (cf. comments on Johnston report). It would be necessary and possible, however, to keep these programs from being either "spy" operations or mutual admiration societies.

The evidence of innovation that appears in these papers is quite encouraging. The concern for evaluation, at least as expressed by these writers, would be encouraging except for the fact that it appears to be going on primarily in ineffable ways. As D. Gordon Rohman suggested in Chapter 6, these new colleges may be most notable for the questions that they have raised.

Now that we find innovation self-consciously pursued, it is time to turn these questions into disciplined inquiry. Campbell's (1969) admonition that we should "shift from the advocacy of a specific reform to the advocacy of the seriousness of the problem, and hence to the advocacy of persistence in alternative reform efforts should the first one fail," should be adopted as a starting point. We should be able to learn from each other's mistakes and successes.[19]

Evaluation reports will benefit by attending more conscientiously to the norms of intellectual inquiry: precise specification of the problem, complete reporting of the method of inquiry and the results obtained, and a statement of conclusions, limitations of the data, and suggestions for further exploration. Narrowing the focus of such reports will probably be distasteful to persons who have had the courage to take on the complex problems of higher education in a holistic manner. Nevertheless, institutional evaluation requires much more explicit reporting than we have yet received. It may be that the best way to put the puzzle together is to pick up one piece at a time.

When evaluation becomes inquiry and embraces both the old and the new colleges, higher education will have become much more socially responsible. When colleges can cooperate in evaluation and when their respective activities, including their classrooms, can be open to external observation, then evaluation can more effectively become inquiry.

References

Campbell, D. T. Reforms as experiments. *American Psychologist,* 1969, **24**, 409-429.

Chickering, A. W., & McCormick, J. Personality development and the college experience. Mimeographed. Plainfield, Vt.: Project on Student Development, Goddard College, 1970.

Suczek, R. F., & Alfert, E. Personality development in two different educational atmospheres. Washington, D.C.: U.S. Office of Education, Bureau of Research, August 1970, Chapter VIII.

---

[19] A fine discussion of an approach to evaluation that effectively combines standardized instruments with a strong reliance on participant observation is the Suczek and Alfert work (1970). Another is Malcolm Parlett's "Evaluating Innovations in Teaching," in *Higher Education in the Seventies,* H. J. Butcher and E. Rudd (forthcoming); and "Undergraduate Teaching Observed," by the same author in *Nature,* Vol. 223, No. 5211, September 13, 1969, pp. 1102-1104.

Chapter 12

EVALUATIVE RESEARCH AND THE NEW COLLEGES

Alexander W. Astin

It is difficult to make any general observation about a series of studies as diverse as the ones presented in this volume. Rather than attempting to comment in detail on each chapter, I should like to present a model of evaluation which tries to put many of the specifics of each chapter in a broader and more systematic context. My philosophy of evaluation, which has been set forth in detail in three recent papers (Astin, 1970a, 1970b; Astin & Panos, 1971), is in some ways similar to what has already been said by several of the authors—in particular Paul Dressel and Arthur Chickering. However, since there are also important differences in our respective approaches, I shall attempt to set forth the basic model, and to illustrate certain of its aspects with reference to the specific chapters.

The Nature of Evaluation

Evaluation in higher education involves the collection of information concerning the impact of some aspect of the institution's educational program. While there are many possible uses for such information, I have assumed that *the fundamental purpose of evaluation is to produce information which can be used in educational decision making.* These decisions may be concerned with the continuation, termination, or modification of an existing program, or with the development and possible adoption of some new program. Whatever the particular decision problem might involve, evaluation is most likely to produce useful information if it is based on an understanding of the nature of the educational decision-making process itself.

Educational Decisions

The need for rendering an educational decision implies the existence of two fundamental conditions: some recognized educational *objective* or set of objectives and at least two alternative *means* for accomplishing these objectives. Although the objectives of educational programs are many and varied, the ultimate objectives of most colleges are usually concerned with the student's learning and personal development, the professor's scholarly work, or the general cultural development of the college community. For the purposes of this analysis, however, I shall direct the discussion to the first of these concerns—the development of the student.

Any educational decision involves a choice among the available alternative means by which the desired objectives may be achieved. In an undergraduate liberal arts college, for example, these means might include the deliberate organization of certain learning experiences (e.g., curricula, instructional methods), the manipulative structuring of the physical environment (e.g., design and location of classrooms, buildings, playgrounds), or the establishment of certain rules and regulations concerning student conduct. Viewed in this way, ***every administrative decision is predicated on a belief in the existence of a causal relationship between some educational objective and a particular means selected to achieve that objective.*** In short, the administrator (or whoever has final responsibility for rendering the decision) believes that of all the means available, the one selected is "best" in the sense that it is most likely to result in the desired outcome.

A Specific Example

Although these assumptions concerning means-ends relationships are often not made explicit, they nevertheless underlie every administrative decision. Take, for example, the apparently minor question of where to locate a new dormitory on the campus. If several alternative sites are available, a decision must be made about which one to use. If the administrator decides to locate the building on site A rather than on sites B or C, his choice is obviously not a random one (although others on his faculty or staff may regard it as such). His decision to build on site A is based on the assumption that, with respect to some outcome, the consequences of building on site A will be superior to the consequences of building on the other sites. The outcome may be an educational one (he believes that the ultimate benefits to the students will be greater if the dormitory is located on site A), an intermediate economic one (the total costs of constructing and maintaining the dormitory will be less), or some combination. Such economic criteria become educational only if the administrator is willing to assume that the student's educational development will be improved in some way by trading off an economy at one point for an increased expenditure somewhere else. In short, the decision to build on site A is necessarily based on an assumption that the chosen alternative (A), in relation to the other possible alternatives (B, C, etc.), is more likely to bring about certain student outcomes that are judged to have value in terms of the goals of the educational program.

Rational decisions in higher education are typically rendered by consulting the available information that is assumed to be relevant both to the desired student outcomes and to the various means under consideration. The principal function of evaluation is to extend this fund of information, in order to enable the decision maker to anticipate better the consequences of the alternative means that he is considering. In the next section I shall discuss the nature of collegiate programs, and then examine some of the evaluative procedures that can be used to produce information for use in making decisions concerning these programs.

The Nature of the College

Any college can be conceived as comprising three conceptually distinct components: *student outputs, student inputs,* and the college *environment.* To be maximally useful in decision making, evaluation in higher education should provide information concerning all three of these program components and their interrelationships.

Student Outputs

Student *outputs* refer to those aspects of the student's development that the college either influences or attempts to influence. Although these outputs can be expressed at very high levels of abstraction (for example, "... the ultimate happiness and well-being of the individual"), we shall limit our consideration of the problem to those relatively immediate outputs that can be operationalized. Specifically, then, student *outputs* would include measures of the student's achievements, knowledge, skills, values, attitudes, aspirations, interests, daily activities, and contributions to society. Other terms that are sometimes used to refer to student outputs are dependent variables, criterion variables, outcome variables, and educational objectives.

Student *inputs* are the talents, skills, aspirations, and other potentials for growth and learning that the new student brings with him to college. These inputs are, in a sense, the raw materials with which the institution has to work. Some inputs may be simply "pretests" on certain student outputs (scores on college admissions tests, for example), while others (sex and race, for example) may be relatively static personal attributes. Inputs can affect the outputs either directly or through interaction with environmental variables.

The *college environment* refers to any aspect of the college or its program that is capable of affecting student outputs. Broadly speaking, the term includes variables such as administrative policies and practices, curriculum, faculty, physical plant and facilities, teaching practices, peer associations, and other attributes of the college experience that might affect the student's development. These environmental variables can, presumably, be changed or manipulated by means of educational decisions.

The relationships among the three components of the model are shown schematically in Figure 1. Note that student outputs can be affected by both environmental variables (Arrow B) and student input variables (Arrow C). Moreover, as Arrow A indicates, college environments can be affected by the kinds of students who enroll. In addition to these "main" effects of environments and inputs on outputs, there may be *interaction* effects involving student inputs and college environments. As the diagram suggests, there are two types of interaction effects: those in which the effect of input on output is different in different environments (AC), and those in which the effect of the environment is different for different types of students (AB).

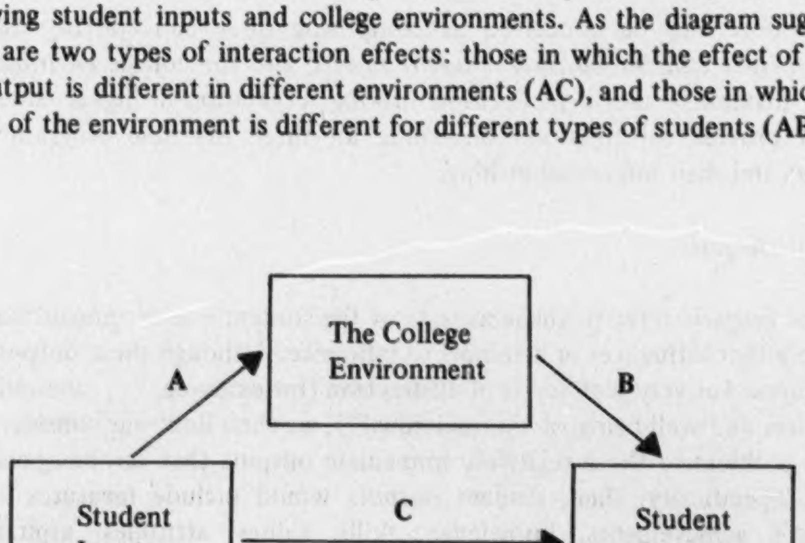


Fig. 1. The three components of the evaluative model.

Different Evaluative Designs

Ideally, any evaluative study involves the collection and analysis of information regarding inputs, outputs, and environments. While the realities of the decision-making process in higher education frequently limit the quality and amount of such information that can reasonably be collected, the usefulness of many evaluative studies could be greatly increased by recognizing more fully the importance of adequate design. In this section I shall review some of the methodologies typically used in evaluative studies, with particular attention to the relevance and possible usefulness of the obtained results in decision making.

Method I: A Description of a Single Educational Environment

This is perhaps the least complex and most widely used method of evaluation. It consists essentially of a logical analysis or simply a detailed description of the educational program being evaluated. A good example of this method would be the procedures typically followed by the regional accrediting associations for colleges and universities. The accreditation process usually involves the collection of descriptive information concerning institutional characteristics, such as faculty-student ratios, teaching loads, size of the library, physical plant, required and elective courses, and the percentage of PhD's on the faculty. Occasionally accreditation may also involve the collection of data on student inputs (e.g, the aptitude test scores or high school grades of the students who are admitted) or student outputs (e.g, the percentage of dropouts or the percentage of students going on to graduate or professional school), although descriptive information on the nature of certain educational operations provides the principal basis for the evaluation.

Another example of a large-scale evaluation using this method is the American Council on Education's survey of the "quality" of graduate programs in various academic disciplines (Cartter, 1966; Roose & Anderson, 1970). Department chairmen and distinguished scholars were asked to rate the quality of the graduate programs offered by each institution granting doctorates in their fields. The resulting average ratings of the "quality" of graduate programs were shown to be highly related to the scholarly productivity and reputation of the faculty in each department.

A descriptive evaluation may involve the use of objective instruments specifically designed to measure characteristics of the college environment, or it may involve simply a narrative or anecdotal description of what the program is like.

Of the several chapters which utilize this approach to evaluation, Chapter 10 on James Madison College is probably the best example. Very detailed information is provided on the philosophy of the program, curricular requirements and options, supportive organizations, administrative structure, physical facilities, communications, and environmental characteristics. Although some data on student inputs are also presented, it is apparently intended to supplement the description of the program rather than to be viewed as a baseline against which output data are to be assessed.

The principal limitation of this method of evaluation is that the descriptive information that is produced is not, by itself, useful in decision making, because it provides no direct evidence about the impact or effect of the various program elements on relevant student outputs. Thus, in the absence of such causal information, the decision maker must *supply* it by assumption. The use of this method to evaluate a curriculum, for example, makes it necessary to assume that "what is taught is what is learned." In evaluating various alternative pedagogical techniques, use of this method ordinarily forces the decision maker to rely on the folklore (e.g., the supposed educational value of seminars versus lectures, or of independent study versus required assignments) in order to make use of the descriptive information about the program.

The studies of graduate education cited earlier provide an interesting example of how the decision maker must supply his own assumptions about the cause and effect in order to use such descriptive information in choosing among the available alternatives. The decision maker, who in this case might be a prospective graduate student in, say, philosophy, wishes to choose the institution where he is likely to get the best possible training in philosophy. If he goes on the basis of the ACE ratings and limits his choice only to those departments that are rated as "outstanding," he is implicitly assuming that the relative educational benefits of different graduate programs in philosophy are proportional to the ACE ratings received by each program. Note that the ratings are not based on any longitudinal information about program impact on students; they are simply the aggregated opinions of departmental quality made by scholars in the field.

If it were possible to supplant the current folklore concerning the effects on educational practices or programs with comparative longitudinal information on student change or growth (e.g., if it could be shown that the graduate student's learning and professional development is actually enhanced more if he attends an "outstanding" institution rather than an "adequate plus" institution), decision makers would have a much sounder basis for using the descriptive information obtained by this method of evaluation.

Method II: Measurement of Students' Outputs

Many measurement specialists are inclined to define evaluation almost exclusively in terms of the assessment of educational outcomes. Typical output measures might include the percentage of dropouts, the percentage of students who subsequently go on to complete the next higher level of education successfully, or the seniors' average performance on standardized tests, such as the GRE.

An advantage of evaluation studies using this method is that they focus attention on the fundamental problems of defining and measuring those outputs that are relevant to the goals of the educational program. Its major weakness is that the obtained output information does not necessarily show how the student has been affected by the particular educational program being studied. Since there is no information bearing directly on the *relationship* between the college's program and the output information, the decision maker is once again forced to supply such causal information by assumption. Typically, this amounts to assuming that what is being measured is the direct result of what has happened to the student in the program.

Student ratings or testimonials about their collegiate experience represent a variation of this method which appears in several chapters. Such ratings are especially attractive to the decision maker, since the student himself is supplying the causal inferences; i.e., the decision maker needs only to "believe" what the student is telling him about how he has been affected by the college's program. Again, the major difficulty here is in determining the extent to which the student's degree of satisfaction with his college is directly attributable to the particular program of the college, and whether he would be any more or less satisfied in a different type of program.

Methods III and IV (see later) both represent versions of what Scriven (1967) has called "noncomparative" evaluation. My general view of the nature of evaluative research is that the impact of any educational practice or program cannot be assessed without resort to comparison with some alternative practice or program. I prefer this "comparative" concept of evaluation for two closely related reasons. First, even in the classical randomized experiment with experimental and control groups, the "control" subjects are not placed either in cold storage or in a state of suspended animation; rather, they are simultaneously exposed to a variety of (often poorly defined) environmental experiences (i.e., if the student were not attending one college, he would be in some other college, or working, or raising a family, etc.). Second, since educational decision making necessarily involves a choice between *alternatives,* and if the decision concerns a possible change in existing practices or the possible adoption of some new program, then the "alternative" which is most analogous to the "control group" is to maintain the status quo.

Method III: Measurement of Environments and Student Outputs

In this method of evaluation, information is collected concerning the educational environments, student outputs, and their interrelationships from two or more colleges or college programs. A simple application of this type of evaluation would be to compare the dropout rate of an experimental college with that of the larger parent university. Any difference would presumably be attributed to differences in the two types of programs. Another typical example would involve comparing the effects of two different methods of teaching a particular course on the student's achievement. The criterion measure in this case might be that student's scores on some standardized achievement test administered after completion of the course; the "measurement" of the educational environment would consist simply of a dichotomy–method A versus method B. A more extensive measurement of the environment would require using several classes which could be ordered systematically (from highest to lowest) on those environmental attributes that may be of interest (e.g., class size, amount of lecture versus discussion, age of instructor, percent of time devoted to independent study, amount of assigned reading, frequency of examination, and so on). Although this latter type of refinement is more cumbersome and expensive, it has the advantage of facilitating the *interpretation* of any observed effects in terms of specific environmental variables. Such information is, of course, much more directly useful to the administrators or faculty who must decide on such matters as class size, assignments, examinations, and so forth.

One advantage of evaluative studies utilizing this design is that empirical data are obtained concerning the relationship between relevant student outputs and program characteristics. The principal limitation of this design is that no information about student inputs is collected. Thus, *unless the students have been randomly assigned to the various educational programs* (which is almost never the case), it is virtually impossible to interpret unambiguously any observed output differences between the programs being compared. For example, does the experimental college have a lower dropout rate than the larger university because of something in its program which captures the student's interest and motivates him to stay in college, or does the experimental college simply attract students who are less dropout prone? Is it possible that the dropout rate of the larger university might even be *lower* than that of the experimental college if it enrolled a comparable group of students? In the absence of any data on differential student inputs to the two programs, there is no way to resolve such ambiguities. The resulting information, as a consequence, may provide a very misleading basis for decision making.

Method IV: Measurement of Student Inputs and Outputs

This method of evaluation involves the measurement of student inputs and outputs. Characteristically, the students complete an attitudinal questionnaire or inventory when they first enter college and take it again one year later, four years later, or in a few cases, many years after graduation. Measures of "change" or "growth" are obtained by comparing the student's input scores from the initial administration with his output scores from the follow-up administration. (These comparative measures usually are simple difference scores, although residual gain scores are used occasionally.) In subsequently interpreting these scores, the investigator typically assumes that any observed changes are due to the students' experiences in college. In other words, he equates "change" with "impact."

Since most of the chapters are concerned with evaluation at only a single college, those that involve collecting longitudinal data can probably be characterized as employing this type of design. The design has the advantage of focusing attention on the longitudinal nature of student change and development in that it views the student's output performance in relation to his input characteristics. Its major weakness, however, is that it really produces no information that bears directly on the question of environmental *impact.* Would the same changes have occured if the student had attended a different kind of college or had not gone to college at all? Since there is only a single environment in such studies, the college environment is not a variable but a constant. (The situation here is identical to the one encountered in experimentation when no control group is used.) Consequently, decision makers who wish to use such information are forced to assume that the same changes would not have occured if the students had attended a different type of college.

The very practical danger in assuming that "change equals impact" can be illustrated with an anecdote. A colleague from a highly innovative small college recently complained to me that nearly a third of his undergraduates who start out majoring in science shift to a nonscience field before graduation. He interpreted this decline in science interest (change) as somehow resulting from the science curriculum of the college (impact). As a consequence, he and other members of a committee on curriculum reform were seriously considering major changes in the science curriculum of the college in the hope of reducing the number of students who withdraw from science fields. As it happened, this colleague's institution was one of several hundred colleges participating in a longitudinal study of institutional impact on career choice (Astin & Panos, 1969). What he did not know was that the longitudinal analyses had revealed that the dropout rate from science actually was *lower* at his college than at almost any other college in the sample. Thus, his college was exerting a relatively *positive* rather than a negative influence on the students' interest in science. Under these circumstances, major changes in the existing science curriculum very well could increase rather than decrease the student dropout rate from science at the college.

The hazards in assuming that change is equivalent to "college impact" suggest that changes in students during college should be viewed as comprising two components: change resulting from the impact of the college and change resulting from other influences (maturation, noncollege environmental effects, etc.). Note that the college may (a) bring about changes which otherwise would not occur, (b) exaggerate or accelerate changes resulting from other sources, or (c) impede or counteract changes resulting from other sources (as in the example, cited above, where the college's dropout rate from science was much lower than average).

Measures of Environments, Student Inputs, and Student Outputs

The ideal design of an evaluative study would incorporate longitudinal data on student inputs and outputs as well as environmental data on at least two contrasting environmental situations (the one which is of primary interest to the decision maker or evaluator and at least one other different environment). This design permits the evaluator to study the growth and development of his students over time, and to determine whether similar changes also occur in students who are confronted with a different kind of environmental experience. Perhaps the only pure example of this type of research in the current volume is contained in the chapter on Michigan's Residential College, where longitudinal changes in students attending the residential college are being compared with changes in students attending the larger parent college of literature, science, and the arts.

One of the practical difficulties with this design is that it requires the evaluator to wait until he has acquired longitudinal data. In Chapter 2 on Research for Action, Arthur Chickering points out that the waiting period required for longitudinal data ordinarily limits the use of such data for decision making. One obvious way to avoid or at least minimize this problem is to collect longitudinal data *routinely,* so that, within a few years, enough longitudinal data will be available to perform any number of "instant longitudinal studies." The routine monitoring of student progress thus provides the decision maker with ready access to longitudinal data from which he can determine how his students are developing and progressing. Such a data system clearly represents a considerable improvement over the typical evaluative study, but it also requires a great deal of advance planning. One possible limitation with this approach is that it may not be possible to anticipate all of the relevant student outcomes at the beginning of data collection, so that information on certain important student inputs may be missing. Again the need for careful advance planning and a clear understanding of the objectives of the program are important in order to reduce the chances that important input information will be neglected.

Another major difficulty with studies using this approach is that it is usually difficult to collect comparable longitudinal student input and output information on colleges other than one's own. Ideally, the evaluation of a new college would involve parallel longitudinal data collection not only at the parent institution (if there is one) and its various other subcomponents, but also on other institutions which may themselves have new colleges. Interinstitutional studies of this type permit the evaluator to assess the importance of the larger institutional context in which his new college is functioning. It would be very informative, for example, to compare the residential colleges at Michigan, Michigan State, the University of the Pacific, and other universities. One suspects from reading over the chapters on these new colleges that the larger institutional context in which each one operates is quite different, even though the new colleges themselves might share many common philosophies and approaches to undergraduate education.

Even if one cannot obtain comparable data from other institutions or even from the parent university because of the high costs or other logistical factors, it is still possible to utilize this design at a more micro level. Rather than comparing longitudinal student data from different institutions, the evaluator can contrast the various subcomponents within his own new college. Such studies might involve a comparison of the effects of peer groups in different residencies, of different curricular emphases, of special programs such as the year abroad versus a traditional year in residency, of different types of financial aid, of varying approaches to counseling and guidance, or even of specific professors' impact as advisers or as teachers. The variety of such "within-college" environmental comparisons is almost limitless. What is needed, however, is that *variability* and *experimentation* be deliberately introduced into the new college, so that contrasting approaches can be compared. Obviously, if the new colleges takes a strong "party line" approach to undergraduate education–regardless of what the party line happens to be–the resulting lack of variability in approach from student to student will make it difficult to identify meaningful subprograms within the college that can be contrasted.

Another potentially serious problem in using this design concerns the statistical methodology for measuring "change" and for contrasting environmental effects under conditions where students entering the different environments may not be comparable at the input stage. Since these problems have been discussed at length elsewhere (Astin, 1970a, 1970b; Astin & Panos, 1971), there is no point in going into these problems in detail. Suffice it to say that it would be wise for evaluators who are not sophisticated in statistical analysis of data to obtain expert consultation before undertaking studies of this type. Even the most sophisticated statistical analysis of such data, however, can and should be made understandable to even the most naive decision maker. The important point for the evaluator and statistician to keep in mind is that the variables must relate to the decision problems which confront the decision maker.

In short, this design has the potential to provide the most useful information to decision makers, since it is directly concerned with how particular educational programs affect the student's development. To make such information maximally useful to decision makers, it is important that the student outputs and the environmental programs that are being measured be relevant to the decision maker's concerns. No matter how elaborate the longitudinal student data and environmental data, and no matter how elegant the statistical analyses, the results will be of only tangential interest to the decision maker if the output variables and environmental variables are not relevant to his particular decision problems.

Evaluation and Innovation

In certain respects it is unfortunate that the need for evaluation has often been associated with educational innovation. Apparently, this association has occurred because of a concern that the enthusiasm of innovators will make it difficult to know whether a new program or approach has really been effective. Thus, one can be encouraged to innovate, but he must remember that he must also "evaluate."

What puzzles me about this line of reasoning is why the status quo is not also put in the same evaluative test. Apparently, one can avoid having his program evaluated as long as he does not change it. But as soon as he proposes changes, particularly drastic ones, there are those who maintain that he must also collect data to demonstrate that the changes are "effective."

This view of evaluation grossly distorts the function of evaluation. The fundamental purpose of any evaluative study, no matter what design the evaluator might use, is to determine the impact of particular educational programs and approaches. The distinction between old programs and new programs is neither necessary nor legitimate. Even in the so-called "formative" evaluations, a given program is contrasted in terms of its various stages of development: the current state is the "new" program and the prior state the "old" program. The current state of the program becomes the "old" program as soon as modifications are made. These changes may occur so rapidly that it would not be possible to compare each stage by means of longitudinal student data. The evaluator must recognize, however, that he is *implicitly* making causal inferences about the relative impact of various stages every time he decides to move from one stage to the next. Thus, "formative" evaluation is not, in principle, different from "summative" evaluation, even though the possiblity of using longitudinal data on student change as a basis for making each decision in the formative stages may be precluded by the rapidity of change. This does not, of course, mean that the evaluator should not collect longitudinal data under such conditions; it means only that he will be limited in the extent to which he can break down his data into units of time that are small enough to enable him to compare separately the relative impact of each formative stage in the development of the new program.

In summary, while it is possible to see how the innovator, as D. Gordon Rohman states, "often looks at the evaluator as the servant of the status quo," there is no necessary reason why evaluation should serve this role. In theory at least, evaluation is just as concerned with the efficacy (or lack thereof) of the status quo as it is with any possible alternatives to the status quo that the decision maker may be considering. In fact, one could regard a commitment to longitudinal evaluation as *requiring* that innovation and experimentation be used routinely in any new college. If a new college takes a monolithic approach to its job of education, there is nothing to evaluate because there is no variability in approach. But if the college deliberately and systematically tries out several contrasting approaches, the possibility of a meaningful comparative evaluative study exists.

Summary

In this chapter I have attempted to present an approach to evaluation in higher education which is designed to produce information which will be maximally useful in decision making. Assuming that the decision maker is primarily concerned with choosing those alternatives that are most likely to enhance the development of the student, there are three different types of data which are necessary in order to make evaluative information relevant to the decision process: student output data, student input data, and data on the college environment. Evaluative studies can be judged in terms of the extent to which they involve all three types of information, and the manner in which the information is analyzed for use in decision making.

Many past and current evaluative studies produce ambiguous findings because at least one of the three informational components is missing. The single-institution or single-program study, through input and output information, indicates how the student changes during college, but it provides no information bearing directly on environmental impact. The multiinstitution or multiprogram cross-sectional study provides information on the relationship between different environments and different outputs, but it is highly susceptible to misinterpretation unless student input data are also collected.

The ideal evaluative study involves multiinstitution or multiprogram longitudinal information on the student inputs, student outputs, and environmental characteristics of the programs being compared.

Evaluation has frequently been misrepresented as something that one must do whenever he wishes to try out innovative programs. This misconception has led some educators to see evaluation as the champion of the status quo and as an obstacle to change. Evaluation does not, however, make any necessary distinction between traditional programs or new programs. Under these circumstances, a traditional or long-standing program is no less in need of "evaluation" than any proposed innovative program. Indeed, a total institutional commitment to an ongoing longitudinal evaluation neccessitates systematic experimentation with new educational approaches in order to provide contrasting environments which can be compared longitudinally.

References

Astin, A. W. The methodology of research on college impact (I). *Sociology of Education*, 1970, 43, 223-254. (a)

Astin, A. W. The methodology of research on college impact (II). *Sociology of Education*, 1970, 43, 437-450. (b)

Astin, A. W., & Panos, R. J. *The educational and vocational development of college students.* Washington, D.C.: American Council on Education, 1969.

Astin, A. W., & Panos, R. J. Evaluation of educational programs. In R. L. Thorndike (Ed.), *Educational measurement.* Washington, D.C.: American Council on Education, 1971.

Cartter, A. M. *An assessment of quality in graduate education.* Washington, D.C.: American Council on Education, 1966.

Roose, K. D., & Anderson, C. J. *A rating of graduate programs.* Washington, D.C.: American Council on Education, 1970.

Scriven, M. The methodology of evaluation. In *Perspective of curriculum evaluation,* AERA Monograph Series on Curriculum Evaluation, No. 1. Chicago: Rand McNally, 1967.

Chapter 13

EVALUATION AND THE NEW COLLEGES: ASSESSMENT FOR STUDENT DEVELOPMENT

Leo A. Munday
Nancy S. Cole

Every generation has to closely examine the social institutions it has inherited to insure that these institutions are serving human needs. Perhaps the disparity between the stated purposes of institutions and later performance exists because institutions change as they age. Or perhaps it is because institutions in our postindustrial society tend to become unwieldy in relating means to specific goals, become impersonal in their operation, and become more concerned with meeting their institutional needs of survival, growth, and material "success" than in meeting human needs. If the "Greening of America" becomes very pervasive, in the next few years more and more people will be asking questions about how the quality of life can be enhanced and what institutions are doing to promote or detract from this quality.

The recent developments of experimental or new colleges can be understood in this context. In part they are attempts at institutional renewal, calling for students, faculty, and administrators to re-think the purposes and corresponding methods of higher education.

Asked to define "new colleges," educators from campus to campus would likely give different emphases reflecting their current concerns and the particular ways their institutions have developed. From the previous chapters, some commonalities can be identified as follows:

1. The new college sets forth new objectives (such as a concern for ecology) or more specific objectives than those generally adopted and then devises a plan to implement the objectives; or

2. The new college holds to the traditional objectives of a college, but reaches for innovative means of implementing the objectives.

3. The new college puts a premium on personalizing education and providing students a freer social environment that allows greater opportunity for impulse expression.

4. The new college develops a campus environment characterized by intellectuality and an academic press.

The Role of Evaluation in the New College

Evaluation has tended to assume a negative role in education, telling educators their new programs are ineffective. Consequently, it is natural for some to consider evaluation the enemy of innovation. However, this view of evaluation is far too narrow. Evaluation in a new college provides fresh opportunities to focus on objectives and how they may be achieved, on students and how they develop in college, and on innovations and how they may be made to work. In each of these three focuses, evaluation may play a more constructive role in a new college than in a traditional college.

New colleges tend to more carefully define their objectives, their goals, and purposes. A familiar problem of evaluation in traditional colleges has been the lack of such carefully specified objectives. Thus, evaluation has often been limited to such criteria of student characteristics as college grades, graduation, or highest degree attained. Undoubtedly, most new colleges would be dissatisfied with these criteria alone (as many traditional colleges are) and they should be. By specifying their objectives in advance in more developmental terms, new colleges have opened the possibility of evaluation based on developmental indicators, on criteria that apply to measurable student growth, at various points in the undergraduate years.

While the focus of all education is the student, the new college has reminded us again how far education can drift away from that focus. By bringing attention back to the student and what happens to him, the new college opens the possibility of evaluation for student development. So much of what passes for evaluation in higher education is for someone other than the student—for the professor, the administration, the governing board, or the people of the state. In the new college, evaluation for the benefit of the student can play a primary role in assuring institutional sensitivity to student needs.

Finally, evaluation may be able to make more constructive contributions in a new college than a traditional college simply because of the innovative nature of the new college. Evaluation becomes an empty act in a situation in which no changes are to be made on the basis of its results. At an innovative new college, eager to find what works, evaluation can be tied directly to action as it ideally should be. Evaluation in an innovative situation leads to truly experimental education.

But it is not enough to note only in the abstract how evaluation may play a constructive role in the new college. To the end of offering specific suggestions, we explore several areas of potential contribution.

Evaluation for Student Development

The breadth of the potential role of evaluation can be seen by following a student from his decision to attend a new college through his years at the college. At the stage of the initial decision, both the college and the student need evaluation information to more adequately consider each other.

Evaluation of a New College for Potential Students

As already noted, new colleges have many different types of goals and objectives. The objectives include student development along nonintellective as well as intellective dimensions. Although colleges are accustomed to communicating their goals for intellectual development to potential students, the nonintellective objectives are less frequently explained. If a new college intends to affect attitude changes in its students, the students should know this before they enroll and take this information into account in their choice of college. Students have a right to know in what ways their attitudes may change as the result of attendance at a particular college, so that they can enter the collegiate situation with more self-awareness and greater responsibility for themselves.

Students are going to make more decisions for themselves, and rely less on institutional decisions made for them. This is shown in college admissions, where students select colleges to a greater extent than colleges select students. In order for students to rationally select, they need information about the differential effects of colleges. While some would say there are no differential effects of colleges, we tend to think that if a college strives to change student attitudes in a certain way, it will more likely do this than a college that does not have this objective. Prospective students need more complete information about the objectives of the colleges they are considering, so they can give more direction to their own development.

Evaluation of Potential Students for the New College

We are accustomed to requiring evaluations of students for college admissions, but we have seen that this is really a reciprocal matter–students also need information to evaluate colleges–and the only question is whether or not choices are to be informed. And, in fact, even the traditional student evaluation requires a broader conception when we consider the new colleges. In line with their purposes, new colleges may desire students with special intellective and nonintellective talents. Thus, multiple student talents must be assessed. With such information, college programs can develop around the characteristics of the students they serve. When a college conceives of its tasks as the development of student talent along the lines the students want to develop, and in ways the college has the expertise to develop, the college will need a wide range of talent indicators available at the time of admissions.

Evaluation in Management of the Environment of a New College

When an innovative educational program is in operation, many decisions must be made. One decides to try this or instead to try that. Thus, at all levels of a new college, evaluative information is essential for proper management. What is going on in the present programs? What trends are developing? Are the right things happening? Systematic collection of evaluative information on a regular basis, say every one-to-three months, can give at least partial answers to these questions–and partial answers are far better than none. Some gross student behaviors which might be considered are: satisfaction of the students with their progress in particular classes, satisfaction of the instructor with what his students are learning, the number of students seeking counseling or seeking to change majors and why, the number leaving college and why, and so on. Administrators and program decision makers should know about such student changes. Just as economists use economic indicators to diagnose, chart, and prescribe for the nation's economic health, perhaps college managers will need comparable indicators to insure that a suitable educational environment develops and continues to respond to students.

Evaluation for Student Growth

We have described the new colleges as student oriented; their primary objectives are to accomplish something with their students. This focus on the student almost requires continuing evaluation of the student's growth and development not for the college only but for the student. It would be helpful to a new college for its instructional and student personnel program to operate within a theoretical framework of how college students develop. Within such a framework, hypotheses could be developed, instruments devised, and appropriate outcomes projected. We could, for example, deal more effectively with such things as the dynamics of peer group influences and their effect on students who are in the minority and majority of students on the campus. Unfortunately, we rarely have such a nice framework. Even so, as we assess college students, the results become more relevant to college needs supportive of basic college purposes if they are related to research on college student development. Thus, local evaluation can benefit considerably from knowledge of the current research in the field.

A basic complication in evaluating student growth is the way students differ from one another in their goals in attending college, as well as in the backgrounds and talents they bring with them. Further, student goals are often quite personal and related to their own self-fulfillment, rather than to the more global goals colleges espouse. This is one reason why it is so hard to study college outcomes and student success, because a certain student behavior (like getting married, getting a job, transferring to another college, or changing majors, to name a few) could represent success for one student and failure to another. This is why we must assess not only to discover how well college goals are being fulfilled but also to help the student discover how well his own personal development is progressing. When the student's own development is assessed, the information can be used by both his college and him to make more rational decisions about what he should study and how he should be taught in order to achieve mutually desired ends. By providing information necessary for wise decisions, evaluations directly promote student growth and development.

Evaluation and the New Colleges–Limitations and Promises

There is no one way to do evaluation. Some people would focus on one aspect and others on another: student output measures, process variables, dimensions of college environment, student characteristics at admissions, and various interactions of these. Some consider growth on cognitive dimensions; others study change along nonintellective lines. Some researchers would let the methodology determine the question to be researched, just as many permit the data available to determine the question to be studied. The answer obtained from evaluation and research in the new colleges is limited only by the creativity of those doing the evaluation. The great promise for evaluation in the new colleges is in asking important questions and basing action on the answers obtained.

Many administrative decisions are made without information. The tragedy is that sometimes information is available but not communicated to the administrator. At other times, some but not definitive information is available and the researcher is reluctant to forward materials he considers to be incomplete. Decisions about program improvement must be based on judgment informed by the best data available, imperfect as they may be. We cannot wait for definitive studies to be done, but a decision made with some information is usually better than one made with no information.

The process by which faculty and administrators design an experimental college can itself be a magic ingredient. Unusual commitment and enthusiasm to a program can have a natural carry-over to students. Perhaps the process of creating a college anew, and involving the principal actors in the creation, is the causative factor in the success of many new colleges.

It is obvious that we cannot expect higher research standards for experimental than for traditional colleges. If evaluation shows that a new program does not get results, it does not follow that we should reject the new in favor of a program which also fails to get results. The evaluation should be aimed constructively at which aspects of a new program are working and in what way changes can be made to improve effectiveness. Evaluators should know only too well the limitations of evaluation, including shortcomings in their methods, instruments, and research strategies, and combine professional humility with a positive attitude toward innovation.

When the possibilities of evaluation of student development are considered, a major gap in our knowledge becomes apparent. In order to reach the objectives of a new college, we must know more about the process of development of the college student. A broader framework of knowledge about how students change during their college years, under what type of influences, and how different influences affect different students is critical to producing and evaluating student growth. Presently, we are guided by incomplete bits and pieces of information. Hopefully, evaluation in the new colleges will contribute to the building of this broad theoretical framework in which a college education can be geared to and evaluated in terms of the growth and development of its students.

What is the new college doing for the "new student" to higher education? Colleges are now receiving more and more students who differ from the traditional college student in these ways: they come from lower socioeconomic status backgrounds, they are the first in their family (or in the first generation) to attend college, they do not have academic attitudes and are oriented to the practical and concrete, and their level of school achievement is more typical of an average high school student than of an average college student. Needless to say, a high proportion are from minority backgrounds. The new colleges described in this book appear more suitable for the traditional college student than for the new student. If this is so, and if the thrust in higher education is for making better provision for the new students, perhaps we should be considering how and along what lines new colleges can be developed for these students. Not only would this field be a socially useful one in which to apply our educational creativity, but the institutions themselves may develop even more stimulating ideas as they grope for programs to serve students with far different needs and backgrounds.

Conclusions

The primary function of all colleges, new and traditional ones, is the fuller educational development of their students. To most of us, this implies student change in cognitive areas and on nonintellective dimensions. Evaluation will be valuable if it does nothing else but remind administrators and faculty that colleges exist for this broad purpose, that student development involves humanistic and intellectual components as well as the familiar ones of socialization and vocational preparation. Beyond budgets, facilities, and committee meetings, colleges serve human needs. And beyond instruments, test scores, and statistics, evaluators help insure that colleges fulfill this service promise.

Chapter 14

THOUGHTS ON EVALUATION AND IMAGINATION

Warren Bryan Martin

What follows is a collection of opinions; opinions about assumptions that may underlie evaluation, about procedures that usually implement it, and about consequences that sometimes result from evaluation. Many of the opinions to be stated here are shared by administrators and faculty in new colleges. Some of these opinions are held by external evaluators of new colleges. A few of the statements to be made find little support beyond the marshland of my own mind.

Despite the probability that ideas contained in this chapter would, taken together, fail to gain widespread support, they do cluster around a widely shared concern. It is the concern, held by program instigators and program evaluators, that there be a proper balance between imagination in launching new colleges and the evaluation necessary to validate them. Everyone is concerned these days for evaluative imagination and imaginative evaluation. Differences begin to emerge when attempts are made to determine what constitutes evaluative imagination and imaginative evaluation. Are there certain ideas being tested in the new colleges, having to do usually with the qualitative dimensions of the teaching and learning relationships, that cannot be evaluated by the methodologies of scientific behaviorism? Could a more humanistic research be devised that would provide a better mechanism for the evaluation of such ideas and their supporting structures and functions? Answers to these questions differ, and contradict each other.

Throughout the essays of this volume, in my judgment, investigators of the new colleges–whether their position is within or outside the institutions–

have too often been willing to measure innovative ideas by conventional research methodologies. Imagination is thus brought to terms with evaluation. I have reacted to this situation by stating, even overstating, the case for other bases for evaluation, differing types of evaluation and, conceivably, no evaluation at all. Also, since the professional evaluators felt free to criticize the reports of the administrators or faculty who presented the evaluations being conducted in the new colleges, I have taken the liberty to evaluate the professional evaluators.

There is a fundamental assumption beneath all of the opinions that are to follow. It is an assumption concerning new colleges, and some lines from T. S. Eliot can be co-opted to make the point:

> Each venture
> Is a new beginning, a raid on the inarticulate
> With shabby equipment always deteriorating
> In the general mess of imprecision of feeling.

New colleges should take their clues from contemporary art forms rather than from the behavioral sciences. This is so because an epoch reveals itself most lucidly in its art, thus giving planners a purchase on existing conditions, and because art, better than science, expresses developing attitudes of men toward man. What men think about man, his nature, cultures, and societies, determines all else–including educational philosophies, curriculum patterns, and governance configurations of new colleges.

The business of the behavioral sciences may be man, but their work is usually descriptive, not normative. What is needed these days is a new vision of man, motivating purposes as expressions of human aspiration, ideas more than analyses. Only in the art forms does insight count as much as sight, and feeling even more than facts. Artists, poets, and musicians, therefore, are better guides to the future than scientists–if we posit a future different than the past. Scientists may extrapolate trends from present realities, but artists provide the best hope for depicting bends in those trends, or what lies beyond them.

The price of being guided by art forms rather than by the behavioral sciences is the necessity to live with Eliot's general mess of imprecision of feeling. And that is a cost which most trustees, administrators, faculty, and students are unwilling to pay. Hence, evaluation, and particularly that which conforms to the methodologies of the behavioral scientists or the systems analysis people.

Evaluators live by the intellect and its instruments, not in the "rag and bone shop of the heart" (Yeats). They favor statistics more than impressions; coefficients of correlation, random samples, and unobtrusive measures, more than immersion in human experience. Experimental colleges are, at best, raids on the inarticulate. Assuming again that the dogmas of the present are inappropriate for the future, the behavioral scientist or systems analyst bring to the evaluative task what can only be described as shabby equipment.

Programs struggling to achieve a future orientation should not be judged by persons whose methodology requires concentration on a present measured by norms from the past. New colleges, therefore, deserve a better fate than they are likely to get at the hands of professional evaluators. In swiftly changing times, research instrumentation cannot keep up with innovative activities, samples will not stand still long enough to provide specimen, contexts are not compatible, hypotheses contradict, methodologies break down.

A moratorium ought to be declared on appraisals of new colleges. After all, appraisal means to put a value on, to evaluate the worth, significance, or status of, especially in terms of an expert judgment. But, alas, colleges are at a time when old values cannot satisfy new conditions, when alleged experts are ill-equipped to measure the worth, significance, or status of that which they would appraise.

Educational research could be a crucial instrument in the determination of institutional priorities and management objectives. In fact, however, it is likely to be of minimal usefulness in these areas. Research should be both purposeful and reflective, that is, for a purpose and from a perspective, but those institutions which research serves are now dissatisfied with past practices and ambiguous about existing purposes, while devoid of perspectives appropriate for the future. Hence, research which is never an end in itself is without an ideational tradition on which it can depend, without a valid context within which to work, and without guidelines by which to plan ahead.

Today, neither theoreticians nor practitioners can speak with confidence about the assumptions or values that should motivate and be the sources of measurement for educational programs. Researchers content themselves, instead, with "behavioral objectives," i.e., quantifiable details, reviews of existing practices, designs for procedural innovations. They are like churchmen who, after giving up on ever gaining agreement on Christian doctrine, decide that they can at least worship together, hoping that preachers who pray together will stay together, no matter what they say or how, or to whom their prayers are directed. A shared liturgy gives the appearance of legitimacy. Anyway, better to light a candle than to curse the darkness.

But that cliche has a fatal flaw. It is not better to light a candle to dispel the darkness when a deadly enemy is lurking in the shadow. To illuminate your candle, if you are the target, is a sure way to get your candle blown out–or your head blown off. Today, research as evaluation, together with all of education, has such an enemy. It is not an interest group or some conspiratorial subversive element, but rather it is the criterion demon.

In medieval Europe the pregnancy of a nun would sometimes be attributed to the devil's demon who had hidden himself in a head of lettuce she ate. Today, when there is no agreement on educational philosophy, when standard tests and measurements have been shown to contain class or race biases, when grading and testing procedures change as often and vary as much as clothing fashions, when there is no agreement on methods of teaching or means of learning (except perhaps that there is no best way), when the primacy of rationality, the "data fixation" of research methodology, the authority of professionalism, the terms of tenure, the meaning of academic freedom, the place and representativeness of governance, the paradigms for student behavior–when all the traditional benchmarks are disputed if not abandoned, what chance is there for research to do more than confirm biases or prejudices previously agreed upon? As researchers peel away layers of institutional lettuce, they find, as nuns suffering from "that pleasing punishment women bear," that the cause of their condition is the criterion demon. But explanations of their plight, and the tactics they use to conceal their condition, are hardly better than those of their medieval precursors.

Theories about man and society, about the purposes of education and the ways it should be organized, about styles of teaching and ways of learning undergird all research efforts. The fundamental assumptions or ideational criteria that motivate and provide the basis for judgments in research continue to be those of straight-arrow western liberalism: confidence in the abilities, worth, and essential goodness of man; preference for individualism rather than community (in this volume–"Thou shalt not interfere with what anyone wants to do."); lip service to democratic processes, to tolerance for diversity, to social pluralism; and commitment to cognitive rationality, coming to the emotions by way of the mind. These are among the intellectual assumptions of educational institutions, academics, and researchers. Furthermore, themes of human development, organization theory, or educational purposes are usually unexamined by researchers even as they are poorly articulated by theoreticians. Research may illuminate reality, but what is less evident yet equally important is that research illuminates researchers' values as well as those of sponsoring agencies.

Today's dogma requires the repudiation of dogma. Concomitantly, old operational assumptions are being exposed and repudiated. As a consequence we are like the foolish professor in that thriller entitled, "We Still Kill the Old Way," who tried to solve a crime in a land where there was no justice. How can evaluation have widespread significance when nobody has criteria that engender widespread trust?

In a period of fiscal stringencies, it is said, new colleges must submit to comparative analysis. Yet, if analyses are made from conventional criteria and with existing instruments, new efforts are doomed–we have met the enemy, and we are theirs. The purpose of the new college is to stand as a place where the old ways not the new ways will have to defend themselves, where the attitudes and actions of participants are not fixed by "those who have gone before," by "those in authority," or even by "impartial observers." To let the competition select the field of battle and determine the rules is to end up playing the game their way–and soon being retired to the sidelines.

If there must be comparative analyses, let them be on terms that will not put the new college at an impossible disadvantage. Evaluators could ask faculty and students in the new unit how they justified the presence of change there, while at the same time asking personnel in the old units how they justify the absence of change there. To the new, "What difference does this innovative program make for people going through it?" And to the old, "What meanings do these conventional programs have for people in them?"

If the inquiry is circumscribed by the methodology of behavioral research, answers will deal with what people do, or perhaps with their attitudes toward what they do. Yet we know that what people say they do seldom conforms with what they actually do. Attention should be given to what respondents do not do, as well as to their attitudes toward what ought to be done. But this gets messy, as messy as life in a time of flux–a condition which is inherently threatening to professional researchers or bureaucratic administrators.

Haven't older programs always undergone evaluation? Didn't the new ones come into existence with a promise to be different and better? Surely then, the new must undergo what the old underwent, in order to prove that it is fulfilling its promise. In most instances, however, older efforts have never been investigated in disciplined ways. What has been learned about them is not usually very important–demographic information and the like. In those few categories where important results have been uncovered, the conclusion of a growing segment of the higher education community is that they don't like what they see–and certainly don't want to emulate it. Yet, to argue that the new must undergo what the old underwent is to imply that the new will be judged by what has been learned from the old, and maybe on its terms. Too bad, especially when, if there is any justification for comparative analyses, it is to determine ways to transform, not merely reform, existing programs.

Can we even be sure of the terms under which the new college began operation? In some cases no particular mandate was given, in others a very general charge launched the venture ("Be different," "Do it better."). Only in a few cases was the new college started with a set of behavioral objectives that yield to empirical evaluation. At several universities where new colleges have been established, the expectations of the original university sponsors changed over time, or new administrators came in with their own commitments and expectations for evaluation. There are few constants against which to measure change. The thinking of the first planners may evolve over time or, leadership attrition being high, their successors may bring other attitudes and interests into play. Therefore, change is the only constant.

After evaluation, policy recommendations will be determined more by the personality, experience, and role of the evaluator than by the substance of the review. Inferences for practice are decided not by research data but by individual values. There is no denying this. Therefore, if comparative analyses are decreed, it is better for innovators to be more concerned with the strategy and tactics of selecting reviewers than with procedures and principles for the review itself. Evaluations are always political, determined more by personal attitudes than by impersonal "facts." So be it. The data were made for man, not man for the data.

Rather than fretting about creating a condition of stability that will allow for a certain type of evaluation, thereby making ends serve means, it might be better not only to tolerate but to actually encourage institutional instability and flexibility. If a new college community is not restless, critical, even dissatisfied with the work of its first planners within five to seven years, ready to say, "the innovations launched here had meaning for earlier participants, but we need experimentation with impact for us," then the college should be judged a failure. We must have places that will make change a habit, not an exception.

Actually, new colleges do make a difference for the people in them. Enough reports from participant observers, case studies, comparative studies, experiential reports, and even some empirical data have been accumulated to make this point.

It is a myth that no evaluation of new programs has been undertaken. Sally Cassidy studied Monteith College thoroughly and appropriately for several years. Jerry Gaff, Paul Heist, and other researchers at the Center for Research and Development in Higher Education have followed cluster college developments. Several of the colleges reported in this book have had respectable evaluation programs that produced useful findings. Those who complain that new colleges are unknown quantities mean just that—the colleges have not yielded to quantitative assessments that, in the judgment of these critics, alone constitute true appraisal. To be sure, in the several studies that have taken place, methodologies varied, variables varied, constraints varied, but somehow, out of all these various efforts, conclusions about the programs have been remarkably similar: The new colleges have effects. Maybe the greater intellectual autonomy, emotional complexity, and sense of community among students and faculty in these programs are not thought to be important or appropriate characteristics. Maybe these changes are due to a self-selective process more than the impact of the environment. Never mind, the point is that something noteworthy happens because these colleges exist.

We know enough about them to conclude that the new ways work as well as, or perhaps better than, the old. Of course no one is satisfied with either what is known, or the extent of the changes perceived. But in a time of general dismay about the rigidities of conventional settings, it is best to err on the side of leniency and patience, for the sake of efforts that venture out on behalf of innovation and experimentation.

New colleges are fragile flowers trying to grow in hard and barren soil and, therefore, they deserve tender care and vigilant protection. Pulling them up periodically by the roots, arbitrarily and with external force, rather than letting them develop organically, is a crime against the future of higher education.

Having stated certain personal biases, I feel disposed to comment on the manifest or latent biases of the professional evaluators. Evaluating evaluators is turnabout fair play. Notice the following assumptions drawn from contributors to this volume, and consider my comments on them:

1. That research will encourage a shift toward reason (Chickering). Comment: From the time of Pierre Bayle (17th century), skeptics have shown that rational effort is its own undoing. Every claim can be countered and has probably been contradicted. To believe that man only acts logically is illogical. Furthermore, as David Hume pointed out, human beings do not proceed from an "is" to an "ought" by logic–or data. The notion that the reasons for research are rational, that research itself is an act of reasoning, that outcomes from research may be expected to encourage the use of reason, is unreasonable.

   Some will say that research doesn't try to proceed from an *is* to an *ought.* Rather, it moves our thinking from "this what seems to be" to "this is what is." Research seeks to describe whatever it inquies into without proceeding into prescription. Research is assessment. Yet, when dealing with realities, the issue of ideals must be raised. In fact, the determination of "what is" will be made in the context of a notion of reality that can only be described as a curious combination of the real and the imagined, the actual and the ideal. Furthermore, research should not ignore policy (surely agitation throughout the last several years about policy research has scored this point), even as it cannot ignore practice. Policy formulation always involves aspiration, hope, and expectation. Thus, be assured that when description passes by, prescription will not be far behind. The systems people, like researchers, usually have the confidence that PMS, MIS, or PPBS will move men toward reason and induce rational behavior. Norman Rae has commented on this tendency in an aphoristic triptych that summarizes my qualifications on such optimism:

   > There are three kinds of Man:
   > Heroic; Foolish; Wise

> The Heroic man continually strives to make order out of chaos; to introduce systems everywhere, even in the face of repeated discouragement.
>
> The Foolish man maintains nothing must be left to chance and is convinced everything can be systematized.
>
> The Wise man knows everything is chance, even the successful working of a system.

2. It is assumed by most researchers that knowledge is power. "The fundamental purpose of evaluation is to produce information which can be used in education decision making [Astin]." Comment: Actually, information may bring paralysis, not power. The problems of the modern intellectual may be caused not by too little formal knowledge, but too much. It is the malady of the cosmic view, what John Barth calls "cosmopsis." Insight, purpose, audacity, these are greater needs; sensitivity, more than knowledge. The purpose of evaluation, then, should be to provide information that will stimulate imagination–not simply gather information.

3. "Evaluation is intrinsic to action [Litten]." Comment: There has been a lot of action for which structured evaluation was not intrinsic, and a lot of action in which evaluation was intrinsic has turned out to be no better than that which did not have it. Furthermore, the type of appraisal advocated by Mr. Litten is, in my judgment, inimical to action. To get conditions necessary for the type of evaluation he promotes will stultify action and induce inactivity. His weapons cannot hit a moving target. The victim must stand still–in order to be acted upon.

4. "Innovation is usually not new or truly experimental [Dressel]." Comment: To be sure, it is rare to encounter anything new under the sun, in education or elsewhere. But it does not follow that an activity–albeit one with ample antecedent or precedent–cannot strike a particular institution or an element in it with the weight of newness. For these places or people, at that time, the old/new change becomes important. Also, an innovation may not be new in concept, but it can be new as an action. Plus, the translation from idea to fact may be new.

5. The present clamor among groups external to institutions of higher education for "accountability" can be answered "with clearly stated objectives and pertinent evidence, the outcomes of sound, evaluative

research [Chickering]." Comment: How can sound evaluative research–whatever that is, and there is no agreement on the definition in this volume–really help answer questions about accountability until there is some measure of agreement concerning that for which educational institutions are to be held accountable? The public or constituency also clamor for efficiency. Can research evaluate efficiency without measures of effectiveness? Do we have them? There are no shared standards for either accountability or efficiency–outside or within the institution. And there will be none until a new cultural rationale or community consciousness is created to replace the one that has collapsed. It is only from such a consensus that operational norms suitable for measuring accountability or efficiency can emerge.

6. Throughout this volume, there are other manifest or latent biases: The "real behavior" approach is one. Techniques are offered whereby, in a variation of the Delphi method, evaluators will be enabled to learn what students, faculty, or others actually do. As acknowledged earlier, since stated attitudes may not square with actions, it is deemed important to chronicle real behavior.

   Three comments: First, today educators dare not settle for reality. We know without research that where we are is not where we ought to be, that what we do is not what we ought to do. There is no need to catalog existing deficiencies. Yet, getting clean data on actual behavior may have the implicit effect of dignifying present conditions. It is easy, therefore, to bring policy into conformity with practice. And that is what we must not do.

   There is another undesirable consequence that can result from this evaluative mechanism. By forcing employment of convergence technique, as with the Delphi method, it may finally be possible to get a diverse group of people to accept certain decisions. But it will be decision making by exhaustion. As with the diminishing food supply in selecting a pope, this is a way to get agreement. But decisions reached under "duress" are suspect.

   A final possibility, of course, is that the camel will be cut and quartered until he can be passed through the eye of the needle. You can accept almost anything if the gestalt of real behavior is broken down into separate units small enough to be handled easily and without threat.

7. Another prevalent assumption is that social forms control individuals. Comment: If this be true, or, to put the point another way, if it is not true

that changed individuals can change institutions, then all we have to look forward to is research organized around program management systems which provide increasingly sophisticated service to the assumptions of the corporate state. The values researchers share will increasingly be those of technology, the organizations that decide their work will continue to be bureaucratic and megatronic. Should the corporate state be the basis for authority, the sanctioner and judge of life?

For this volume, participants in the development of colleges committed to innovation or experimentation have reported on evaluation mechanisms and, in some cases, on outcomes of their evaluations. Additionally, professional evaluators have discussed or commented on the institutional evaluations reported. The volume also includes comments by observers of the processes of change in American higher education.

It is evident that there are differing and often rival interpretations among contributors to the book concerning the intention of evaluation, its most effective means of implementation, how much should be reported, and the nature or purposes of the new colleges themselves.

However frustrating these divergencies may have been to the reader, perhaps it is well for so much variety to exist. Assuming that leaders are uncertain about the course to be taken by postsecondary education in the future, the best thing that can now be done may be to multiply options and exercise them. Different institutions, with differing programs, evaluated differently, should ultimately lead us to understand and accept the infinite variety of educational opportunities required to meet the infinite capabilities of human beings.

Especially needed now are fresh efforts at creative syntheses. On most issues–the interaction of mind and body in learning, behavioristic and humanistic research methodologies, realism or idealism in educational philosophy, external standards and internal motivation, the authority of the person compared with that of the crowd, the advantages of institutional largeness or smallness, the setting, timing, and styles best suited for institutional change, conventional instruction compared with innovative or experimental approaches to it–on these and other issues that divide us, the fact is that all sides are "right." This is not to say that there are no wrongs. (Pluralism has its limits–it cannot be indifferent to order or social justice, for example.) Rather, it is to assert that for most of the problems confounding educators today, including the problem of evaluation, rival theorists and practitioners have essential contributions to make.

Professional evaluators are right in insisting that, despite a paucity of rigorous assessment, traditional programs must be credited with giving youth professional or vocational training (albeit training that is too often dated or otherwise inappropriate), as well as some more general knowledge and sociopolitical sophistication than that which can be expected to come to a person simply through the process of maturation. The college or university experience makes a difference.

Obviously, not all of that experience is positive in its effects, indeed much of it may be negative, and new colleges were brought into existence to provide correctives for present inadequacies. They dare to offer alternatives. But unless such colleges are willing to enunciate institutional commitments, examine their objectives and procedures, or be examined by others, what chance is there for determining whether significant changes are taking place? To devise, revise, and improvise, without knowing why, how, or so what, is to promote confusion and assure disillusionment. The effects may be especially devasting for students, the interest group new colleges are usually most eager to help. The youth can have no way of knowing what they are getting into or what they are getting out of it. The consequences of this condition for an institution will be complete permissiveness or eventual reaction culminating in imposed definitions of probity or rectitude—the resuscitation of that state of affairs which had supposedly been put to rest.

There must be, there will be evaluation. The question is: Evaluation of what sort, for what purpose, to what end? Advocates of alternative schools are insightful in their assertion that behavioristic research and systems analysis may be inappropriate means of evaluating new colleges. Assessment procedures should be congruent with program objectives. Yet, as Paul Dressel has said, evaluation need be neither so complicated or exclusively quantitative as some researchers would make it, nor so devastating or irrelevant as many innovators fear. There can be creative evaluation, responsive to the imagination of new colleges, yet useful in providing comparisons with that which exists elsewhere.

Although we suffer from the tendency to dichotomize (as dramatized in this chapter), too often throwing up our hand at prospects for diversity of pluralism, or, more specifically, at prospects for alternative models and disciplined appraisals of them, bleakly concluding that the American institution of higher education cannot have it both ways, perhaps the greatest challenge now is to guarantee that we do just that—have it both ways. The future belongs to educational institutions able to combine imagination and evaluation. That, in fact, is the only viable future.

THE AMERICAN COLLEGE TESTING PROGRAM

The American College Testing Program conducts its nationwide and international educational service activities from a headquarters complex, six regional offices, four subregional branches, and a Washington, D.C., office. The national headquarters building is located in suburban Iowa City, Iowa, at 2201 North Dodge Street. General mailing address is P.O. Box 168, Iowa City, Iowa 52240.

National Officers

Fred F. Harcleroad, *President*
Oluf M. Davidsen, *Vice President-Program Operations*
Marvin F. Brecht, *Vice President-Business and Finance and Treasurer*
Kenneth E. Young, *Vice President-Washington, D.C. Office*
Arthur E. Smith, *Vice President-Educational Services*
Leo A. Munday, *Vice President-Research and Development*

For Consultation Assistance

Western Regional Office
Downtown Plaza Towers, Suite 515
555 Capitol Mall
Sacramento, California 95814

Mountain-Plains Regional Office
Executive Office Building
720 Pearl Street
Boulder, Colorado 80302

Kansas-Nebraska Office
P.O. Box 1104
Manhattan, Kansas 66502

Southwestern Regional Office
Coronado Building
1626 50th Street
Lubbock, Texas 79412

Texas Office
909 Dalworth, Suite 201A
Grand Prairie, Texas 75050

Oklahoma-Arkansas Office
Nichols Hills Exec. Bldg., Suite 105
6403 N.W. Grand Boulevard
Oklahoma City, Oklahoma 73116

Midwestern Regional Office
899 Skokie Boulevard
Northbrook, Illinois 60062

Ohio-Indiana Office
133 North Prospect Street
Bowling Green, Ohio 43402

Southeastern Regional Office
20 Perimeter Park, Suite 101
Atlanta, Georgia 30341

Eastern Regional Office
General Washington Building
216 Goddard Boulevard
King of Prussia, Pennsylvania 19406

ACT PUBLICATIONS

The ACT publications listed here may be obtained, if copies are available by writing to: ACT Publications, P.O. Box 168, Iowa City, Iowa 52240.

ACT and the Two-Year College: A Guide for Using ACT Services in Two-Year Institutions. Describes ACT services and programs designed specifically for junior and community colleges and vocational-technical institutions. Free.

ACT Financial Aid Services. Describes ACT's three major financial aid services available for use by students and institutions of higher education. Explains operation of systems used in ACT Student Need Analysis Service. Free.

College Student Profiles: Norms for the ACT Assessment. Describes students, by geographic region and institutional level, enrolled in ACT colleges and universities. $3.50, or $3.00 if check accompanies order.

Counselor's Handbook: A Guide for Using ACT Services in Pre-College Guidance. Provides secondary school counselors with pertinent and accessible information for use in advising college-bound students and their parents. Free.

The Developing State Colleges and Universities: Historical Background, Current Status, and Future Plans, by Fred F. Harcleroad, H. Bradley Sagen, & C. Theodore Molen, Jr. Describes a project undertaken to determine the extent and kind of change that has taken place in the developing state colleges and universities in the past and is projected for the near future. $3.00.

Educational Opportunity Service Announcement. (Brochure) Summarizes a new institutional service designed to inform prospective students of special or unusual educational or scholarship programs available at the institution, and outlines procedures necessary for participation. Free.

Institutional Self-Study Service. (Brochure) Describes the purpose, design, and features of the service for assessing the quality of education provided at specific institutions. Free.

Quiz for College Presidents and All Others Who Care about Students. Features a quiz (with answers) on college students, plus a description of the mechanics and philosophy of the overall ACT Program. The seven-part quiz gives an insight into the thinking of today's college students. Free.

Using ACT on the Campus. A guide for the use of ACT services at institutions of higher education. Provides assistance in understanding and using ACT data in colleges for admissions, scholarship selection, course sectioning, and personnel services. Free.

When You Listen, This Is What You Can Hear, by Gordon A. Sabine. The results of YOUTHPOLL, a mail questionnaire survey conducted in the fall of 1970, among a sample of college freshmen and high school seniors who had taken the ACT Battery the preceding year. The poll gave the students the opportunity to comment candidly about their parents, their schools and teachers, and student protest. $3.00.

Monographs

ACT Monographs on selected topics in educational research are published periodically as a public service. Copies may be obtained for $3.00 each, if available, by writing to: ACT Publications, P. O. Box 168, Iowa City, Iowa 52240.

No. 1 *Computer Applications to Financial Aid Processing,* by R. M. Brown, S. S. Janes, & I. O. Overman (Out of print)

No. 2 *The Two-Year College and Its Students: An Empirical Report,* by the staff of the Research and Development Division

A collection of several special reports concerning two-year colleges and their students. The focus is on institutional characteristics, student characteristics, and predictions.

No. 3 *The Ghetto College Student: A Descriptive Essay on College Youth from the Inner City,* by G. D. Morgan

Provides an enlightened picture of young people from the ghettos. It is based on a series of interviews of teachers as well as students.

No. 4 *Open Admissions and Equal Access,* by P. R. Rever, Ed.

Based on selected papers read at the 1970 annual meeting of the American Association for Higher Education.

No. 5 *Financing Higher Education: Alternatives for the Federal Government,* by M. D. Orwig, Ed.

A collection of special papers, divided into four basic parts: (1) The Economic and Social Background; (2) Who Should Pay: Alternative Strategies for Financing; (3) Alternatives for the Federal Government; (4) The Road Ahead.

No. 6 *Assessment in Colleges and Universities,* by Fred F. Harcleroad and Jean H. Cornell, Eds.

Proceedings of the 1970 ACT Invitational Conference.

Research Reports

The ACT Research Reports are published periodically by the Research and Development Division of The American College Testing Program. Copies of the reports may be obtained for $1.00, if available, by writing to the Research and Development Division, The American College Testing Program, P.O. Box 168, Iowa City, Iowa 52240. Check or money order must accompany request.

No. 1 *A Description of American College Freshmen,* by C. Abe, J. L. Holland, S. W. Lutz, & J. M. Richards, Jr.

No. 2 *Academic and Nonacademic Accomplishment: Correlated or Uncorrelated?* by J. L. Holland, & J. M. Richards, Jr.

No. 3 *A Description of College Freshmen: I. Students with Different Choices of Major Field,* by C. Abe, & J. L. Holland

No. 4 *A Description of College Freshmen: II. Students with Different Vocational Choices,* by C. Abe, & J. L. Holland

No. 5 *A Description of Junior Colleges,* by J. M. Richards, Jr., L. M. Rand, & L. P. Rand

No. 6 *Comparative Predictive Validities of the American College Tests and Two Other Scholastic Aptitude Tests,* by L. A. Munday

No. 7 *The Relationship between College Grades and Adult Achievement: A Review of the Literature,* by D. P. Hoyt

No. 8 *A Factor Analysis of Student "Explanations" of Their Choice of a College,* by J. M. Richards, Jr., & J. L. Holland

No. 9 *Regional Differences in Junior Colleges,* by J. M. Richards, Jr., L. P. Rand, & L. M. Rand

No. 10 *Academic Description and Prediction in Junior Colleges,* by D. P. Hoyt, & L. A. Munday

No. 11 *The Assessment of Student Accomplishment in College,* by J. M. Richards, Jr., J. L. Holland, & S. W. Lutz

No. 12 *Academic and Nonacademic Accomplishment in a Representative Sample Taken from a Population of 612,000,* by J. L. Holland, & J. M. Richards, Jr.

No. 13 *The Prediction of Student Accomplishment in College,* by J. M. Richards, Jr., J. L. Holland, & S. W. Lutz

No. 14 *Changes in Self-Ratings and Life Goals among Students at Colleges with Different Characteristics,* by R. W. Skager, J. L. Holland, & L. A. Braskamp

No. 15 *Can Computers Write College Admissions Tests?* by J. M. Richards, Jr.

No. 16 *Changes in Self-Ratings and Life Goals as Related to Student Accomplishment in College,* by R. W. Skager, & L. A. Braskamp

No. 17 *Family Income and the Characteristics of College-Bound Students,* by L. L. Baird

No. 18 *Predicting a Student's Vocational Choice,* by J. L. Holland, & S. W. Lutz

No. 19 *The Educational Goals of College-Bound Youth,* by L. L. Baird

No. 20 *Who Goes Where to Junior College?* by J. M. Richards, Jr., & L. A. Braskamp

No. 21 *Predicting Student Accomplishment in College from the ACT Assessment,* by J. M. Richards, Jr., & S. W. Lutz

No. 22 *The Undecided Student: How Different Is He?* by L. L. Baird

No. 23 *The Effects of Selecting College Students by Various Kinds of High School Achievement,* by L. L. Baird, & J. M. Richards, Jr.

No. 24 *Do They Do What They Say They Will Do?* by S. W. Lutz

No. 25 *Changes in the Vocational Plans of College Students: Orderly or Random?* by J. L. Holland, & D. R. Whitney

No. 26 *The Flow of High School Students to Schools, Colleges, and Jobs,* by L. L. Baird, & J. L. Holland

No. 27 *Forecasting Academic Success in Specific Colleges,* by D. P. Hoyt

No. 28 *A Description of Graduates of Two-Year Colleges,* by L. L. Baird, J. M. Richards, Jr., & L. R. Shevel

No. 29 *An Empirical Occupational Classification Derived from a Theory of Personality and Intended for Practice and Research,* by J. L. Holland, D. R. Whitney, N. S. Cole, & J. M. Richards, Jr.

No. 30 *Differential Validity in the ACT Tests,* by N. S. Cole

No. 31 *Who Is Talented? An Analysis of Achievement,* by C. F. Elton, & R. Shevel

No. 32 *Patterns of Educational Aspiration,* by L. L. Baird

No. 33 *Can Financial Need Analysis Be Simplified?* by M. D. Orwig, & P. K. Jones

No. 34 *Research Strategies in Studying College Impact,* by K. A. Feldman

No. 35 *An Analysis of Spatial Configuration and Its Application to Research in Higher Education,* by N. S. Cole, & J. W. L. Cole

No. 36 *Influence of Financial Need on the Vocational Development of College Students,* by A. R. Vander Well

No. 37 *Practices and Outcomes of Vocational-Technical Education in Technical and Community Colleges,* by T. G. Gartland, & J. F. Carmody

No. 38 *Bayesian Considerations in Educational Information Systems,* by M. R. Novick

No. 39 *Interactive Effects of Achievement Orientation and Teaching Style on Academic Achievement,* by G. Domino

No. 40 *An Analysis of the Structure of Vocational Interests,* by N. S. Cole, & G. R. Hanson

No. 41 *How Do Community College Transfer and Occupational Students Differ?* by E. J. Brue, H. B. Engen, & E. J. Maxey

No. 42 *Applications of Bayesian Methods to the Prediction of Educational Performance,* by M. R. Novick, P. H. Jackson, D. T. Thayer, & N. S Cole

No. 43 *Toward More Equitable Distribution of College Student Aid Funds: Problems in Assessing Student Financial Need,* by M. D. Orwig

No. 44 *Converting Test Data to Counseling Information,* by D. J. Prediger